

VOCABULARIOS ESTRUCTURADOS EN LA WEB

Visualización e integración con las
aplicaciones



VNIVERSIDAD
SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

TRABAJO DE FIN DE GRADO

Realizado por Paula Sánchez García
y dirigido por Carmen Caro Castro

Universidad de Salamanca, 2013

**UNIVERSIDAD DE SALAMANCA
FACULTAD DE TRADUCCIÓN Y DOCUMENTACIÓN
GRADO EN INFORMACIÓN Y DOCUMENTACIÓN
Trabajo de Fin de Grado**

VOCABULARIOS ESTRUCTURADOS EN LA WEB.

Visualización e integración con las aplicaciones.

Autora: Paula Sánchez García

Tutora: Carmen Caro Castro

Salamanca, 2013

ASIENTO CATALOGRÁFICO ADAPTADO AL REPOSITORIO INSTITUCIONAL GREDOS

Título:

Vocabularios estructurados en la web: visualización e integración con las aplicaciones

Autor:

Sánchez García, Paula

Director:

Caro-Castro, Carmen

Palabras clave:

[ES] Tesauros-normas, lenguajes documentales, búsqueda de información, recuperación de información, navegación, visualización, integración

[EN] Thesaurus-standards, indexing languages, information search, information retrieval, browsing, display, integration

Clasificación UNESCO:

57 Lingüística: 5701 Lingüística aplicada: 5701.05 Lenguajes documentales

Fecha:

2013-07-7

Descripción:

Trabajo de Fin de Grado en Información y Documentación, curso 2012-2013

Resumen:

[ES] Desde la aparición de los primeros tesauros, la normativa que regula su gestión y elaboración ha ido evolucionando para adaptarse a los nuevos tiempos y en especial, al entorno digital. Este trabajo se centra en comprobar si los tesauros electrónicos han aprovechado este nuevo entorno y las oportunidades que brinda la Web Semántica para cumplir la última normativa publicada (la ISO 25964) en materia de visualización de la información, búsqueda y navegación. Por último, se analiza si estos instrumentos están integrados con algún tipo de aplicación en la web

Abstract:

[EN] Since the appearance of the first thesaurus, the standards governing their management and development have evolved to adapt to changing times and especially in the digital environment. This paper focuses on checking whether electronic thesauri have taken advantage of this new environment and the opportunities provided by the Semantic Web to meet the latest regulations published (ISO 25964) in the field of information display, searching and browsing. Lastly, it analyzes whether the instruments are integrated with some kind of application on the web.

SUMARIO

Asiento catalográfico adaptado al repositorio institucional GREDOS	I
Sumario	II
Índice de imágenes	IV
Índice de tablas	IV
Índice de gráficos	V
I. INTRODUCCIÓN	1
1. Justificación	1
2. Estructura	2
II. MARCO TEÓRICO	4
3	
1. Definición, origen y evolución de los tesauros	4
1.1 Definición	4
1.2 Origen y evolución	5
2. Evolución de la normativa relativa a tesauros	7
2.1 Primeras normas	7
2.2 Norma ISO 25964	10
2.3 Norma UNE 50-106-90 y UNE 50-125-97	11
3. Visualización y navegación e integración con las aplicaciones	12
3.1 Visualización	13
3.2 Navegación y búsqueda	17
3.3 Integración con las aplicaciones	18
III. ESTUDIO	20
1. Objetivos	20
2. Metodología	20
3. Resultados y análisis	23
3.1 Visualización	23
3.1.1 Tipos de visualización	23
3.1.2 Visualización del registro simple	25
3.1.3 Visualización alfabética	26
3.1.3.1 Características de la visualización alfabética	28
3.1.4 Visualización jerárquica	28
3.1.4.1 Características de la visualización jerárquica	29
3.1.4.2 Información adicional en la jerarquía	29
3.1.5 Visualización gráfica	30
3.1.6 Presentaciones multilingües	31
3.1.6.1 Características de las presentaciones multilingües	31
3.1.7 Gráfico general del apartado de visualizaciones por tesauros	33

3.2 Búsqueda	34
3.3 Navegación	37
3.4 Búsqueda y navegación	39
3.5 Integración con las aplicaciones	40
3.6 Valoración general	42
IV. CONCLUSIONES	44
V. BIBLIOGRAFÍA	47
VI. ANEXOS	52

ÍNDICE DE IMÁGENES

Imagen 1. Evolución de la normativa sobre tesauros (Dextre-Clarke; Lei Zeng, 2011)	8
Imagen 2. Tesauro EUROVOC (1990). Lista alfabética estructurada (fragmento)	14
Imagen 3. Tesauro EUROVOC (en línea). Registro simple	14
Imagen 4. Tesauro EUROVOC (en línea). Visualización sistemática	15
Imagen 5. Tesauro de la UNESCO (en línea). Visualización sistemática	15
Imagen 6. Tesauro EUROVOC (1990). Índice alfabético	16
Imagen 7. Tesauro EUROVOC (en línea). Visualización alfabética	16
Imagen 8. Tesauro EUROVOC (1990). Terminograma	17
Imagen 9. Social Semantic Web Thesaurus. Presentación gráfica	17
Imagen 10. Ejemplo del tesauro de Biblioteconomía del CINDOC con visualización alfabética	26
Imagen 11. Ejemplo del tesauro AGROVOC: búsqueda por “a”	27
Imagen 12. Ejemplo de búsqueda avanzada en el tesauro EUROVOC: búsqueda por “u”	27
Imagen 13. Visualización gráfica del concepto “decision-making” en el tesauro Reegle	31

ÍNDICE DE TABLAS

Tabla 1. Tipos de visualización	24
Tabla 2. Registro simple	25
Tabla 3. Valoración general	43

ÍNDICE DE GRÁFICOS

Gráfico 1. Tipos de visualización	23
Gráfico 2. ¿Cuántos gráficos tienen información adicional?	30
Gráfico 3. ¿Qué información adicional contienen?	32
Gráfico 4. Características de la visualización multilingüe	33
Gráfico. 5. Características de la visualización por tesauros	33
Gráfico 6. Opciones de búsqueda	35
Gráfico 7. Características de la búsqueda por tesauros	36
Gráfico 8. Opciones de navegación	37
Gráfico 9. Características de la navegación por tesauros	38
Gráfico 10. Características de la búsqueda y navegación por tesauros.	39
Gráfico 11. Integración	40
Gráfico 12. Características de la integración por tesauros	41
Gráfico 13. Valoración general	42

I. INTRODUCCIÓN

1. Justificación

Internet llegó, convenció y se quedó. Este conjunto de redes de comunicación interconectadas ha revolucionado todos los ámbitos de la vida diaria de los seres humanos: educación, salud, medios de comunicación, moda, fotografía, música, actividades bancarias, tareas administrativas, etc. Todos y todo ha caído en esta telaraña que ha conseguido envolver el mundo, haciendo que su uso sea casi imprescindible en nuestro día a día.

Internet no solo ha influido en todos esos ámbitos, sino que también ha propiciado la transformación de muchos elementos de nuestra vida cotidiana en un formato electrónico: libros en papel en libros electrónicos, programas de radio a través de las ondas en programas de radio a través de Internet, movimientos bancarios en papel en movimientos bancarios a través de la red, etc. Y por supuesto, también, tesauros impresos en tesauros electrónicos.

Como se observa, ningún campo ha escapado a esta revolución, y la Información y Documentación no es una excepción.

Los tesauros han estado relacionados con la informática desde su aparición, aunque no ha sido hasta finales del siglo XX cuando este hecho ha sido recogido por las recomendaciones y los organismos internacionales. Con esto queda demostrado la vigencia e importancia de estas herramientas en la actualidad (Sánchez-Cuadrado; Colmenero-Ruiz; Moreiro, 2012, p. 234).

Desde su creación, los tesauros, han gozado de una opinión muy positiva entre los profesionales de la Información y la Documentación, y han desarrollado una función muy importante de almacenamiento y recuperación de la misma (Mochón Bezares; Sorli Rojo, 2010, p. 644).

Sin embargo, los instrumentos de la época inicial se han quedado obsoletos. Por eso, actualmente hay un continuo desarrollo de nuevas herramientas que buscan seguir mejorando y solucionando las necesidades de los usuarios.

La necesidad de transformar estas herramientas a formato electrónico fue totalmente inevitable para ayudar a normalizar toda la masa ingente de información que hay en Internet (de igual manera que hasta ese momento los tesauros han ayudado a organizar los contenidos impresos). De esta manera, estas estructuras de información se convirtieron en elementos esenciales también en el ámbito electrónico, y con una amplia difusión.

Además, también es un hecho que muchos tesauros aparecieron directamente en formato electrónico como si fueran nuevos, es decir, sin estar antes en formato impreso, cuando los organismos se dieron cuenta del potencial de Internet.

Esta transformación tan radical, pero tan necesaria a la vez, ha supuesto un cambio muy importante en su normativa y estandarización, propiciando la aparición de multitud de normas que recogen este hecho.

Y en este trabajo tratamos precisamente de eso, de descubrir si los tesauros electrónicos cumplen la normativa en determinados aspectos como son la presentación, la búsqueda, la navegación y la integración con bases de datos u otras aplicaciones. Para ello, nos basaremos en la última norma ISO publicada sobre tesauros, la *ISO 25964 Information and documentation — Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval*, y realizaremos un cuestionario para comprobar el grado de cumplimiento de dicha norma en algunos de ellos.

Elegimos el tema de los tesauros por ser una herramienta básica en nuestro campo de trabajo, y porque para los profesionales que trabajamos en ello es fundamental contar con instrumentos que cumplen la normativa actual y que están actualizados. Además, en la literatura científica existe una escasez (por no decir total carencia) de trabajos de investigación cualitativos con la norma ISO 25964-1 como base que analicen los aspectos concretos que nosotros tratamos en este trabajo sobre los tesauros electrónicos, por lo que se podría decir que es un trabajo innovador. El único trabajo que trata algunos apartados parecidos al nuestro es un estudio realizado por Gonzalo Mochón Bezares y Ángela Sorli Rojo en 2010. En él analizaron multitud de características (entre ellas, los formatos de presentación) de 172 tesauros que se encontraban en acceso abierto en Internet.

Ha sido un trabajo duro y constante a lo largo de un extenso período de tiempo, pero las conclusiones a las que hemos llegado merecen todo el esfuerzo realizado.

2. Estructura

La estructura del trabajo tiene como base la normativa para la creación de Trabajos de Fin de Grado en Información y Documentación, aunque con algunas modificaciones para facilitar su exposición.

La primera parte consta de los apartados básicos de cualquier trabajo de investigación: cubierta, portada, asiento catalográfico, resumen, palabras claves en español e inglés, sumario e introducción. En este último apartado se presenta el trabajo, justificando el tema elegido, y se explica la estructura.

La segunda parte la conforma el marco teórico de la investigación, centrado principalmente en mostrar un contexto general del tema basado en cuatro puntos: definición, origen y evolución de los tesauros, evolución de la normativa relativa a esta herramienta, comparativa de

visualizaciones entre la norma UNE 50-106-90 y la ISO 25964-1, y la integración de estos elementos con otras aplicaciones.

La tercera parte muestra el trabajo en sí, es decir, los objetivos principales y específicos que nos hemos planteado conseguir, la metodología llevada a cabo para alcanzarlos (detallando la justificación de la selección de la muestra, la elección de las variables, el método de análisis empleado, etc.) y los resultados de la evaluación que hemos hecho a los tesauros. Estos resultados se exponen a través de unos gráficos y tablas realizados con Excel con su correspondiente comentario.

En la cuarta parte reflejamos las conclusiones a las que hemos llegado a la luz de los resultados obtenidos y de los objetivos previamente planteados.

La última parte del trabajo recoge la bibliografía utilizada para poder realizarlo y un anexo donde se ha adjuntado el cuestionario que se realizó para evaluar los tesauros.

II. MARCO TEÓRICO

1. Definición, origen y evolución de los tesauros

1.1 Definición

El propósito principal de un tesoro es guiar tanto al indizador como al usuario para que elijan siempre el mismo término preferente (o una combinación de términos preferentes) cuando quieran representar un concepto o un tema determinado. Para lograr este fin, en primer lugar se debe crear una lista de todos los conceptos que podrían ser útiles para representar y recuperar información en un área de conocimiento. Estos conceptos se designan por términos, en ocasiones por más de uno, por lo que habrá que seleccionar un término preferente para representarlos. En segundo lugar, el tesoro debe presentar los términos de tal manera que los usuarios puedan identificar fácilmente los que necesitan. Esto se logra mediante el establecimiento de las relaciones entre términos y/o entre conceptos y presentando estas relaciones de una forma lógica y estructurada (ISO 25964-1, 2011, p. 13).

La mayor parte de los teóricos de la información y la documentación ha desarrollado su propia definición de tesoro. Sin embargo, como este trabajo tiene un enfoque esencialmente normativo, nos centraremos en las definiciones que ofrecen los distintos estándares y finalmente incluiremos nuestra propia definición.

La norma ISO 2788 (1986) los define desde dos puntos de vista:

- a) por su función, como un instrumento de control terminológico utilizado para trasponer a un lenguaje más estricto el idioma natural empleado en los documentos y por los indicadores
- b) por su estructura, como un vocabulario controlado y dinámico de términos que tienen entre ellos relaciones semánticas y genéricas y que se aplica a un dominio particular del conocimiento (ISO 2788, 1986).

La norma UNE 50-106-90 indica que son un vocabulario de un lenguaje de indización controlado, organizado formalmente con objeto de hacer explícitas las relaciones entre conceptos (UNE 50106, 1990, p. 5).

La norma ISO 25964-1 explica que se trata de un conjunto de conceptos seleccionados que están representados principalmente por términos preferentes organizados formalmente para que las relaciones paradigmáticas entre los conceptos se hagan explícitas, y los términos preferentes se acompañen de sinónimos o cuasi sinónimos (ISO 25964-1, 2011, p. 10).

En conclusión, podemos definir el tesauro como un conjunto de términos normalizados que sirven para representar conceptos o temas del lenguaje natural y en los que quedan patentes las diversas relaciones (asociativa, jerárquica, etc.) entre los mismos.

1.2 Origen y evolución

El primer origen del término “tesauro” está en el griego y su significado era el de “colección”. El latín lo absorbió y le dio el significado de “tesoro”. Su primer uso moderno data de 1531 aunque la palabra fue usada por primera vez por Latini (1220-1294) en su *Livre du Tresor*, una enciclopedia que fue escrita entre 1262 y 1264 en francés (Gil Urdiciain, 1998a, p. 64).

A mediados del siglo XIX, concretamente en 1848, Peter Mark Roget (1779-1869) publicó el *Thesaurus of English words and phrases classified and arranged so as to facilitate the expression of ideas and assists in literary composition*. Era una clasificación ideológica de un vocabulario donde relacionaba por sinonimia los términos más importantes de una lengua. Aunque su fin era todavía más lingüístico que documental, su rasgo de identidad fue la manera de presentar los términos: a través de una cadena de sinónimos que formaban categorías conceptuales. Primaba el significado sobre la ortografía (las anteriores obras sobre sinónimos se basaban en una presentación alfabética). Esta publicación fue fundamental para transformar ideas en palabras y se puede considerar el origen de los tesauros documentales

Como se lee, en los primeros tiempos el tesauro se definía como un vocabulario no especializado pero restringido a una lengua, por lo que tenía un uso puramente lingüístico.

Los primeros tesauros “documentales” (Lancaster, 1995) aparecieron como consecuencia de la necesidad de crear instrumentos de control terminológico que facilitaran la representación y recuperación de una información que cada vez era más especializada. (López Huertas, 1991)

La primera persona que utilizó el término tesauro en este sentido fue Helen Brown cuando escribió en 1957 (Dorking Conference on Classification): "El problema de la recuperación de información es transformar conceptos y sus relaciones, como se expresan en los documentos, en un lenguaje más regularizado, con los sinónimos controlados y sus estructuras sintácticas simplificadas". Se mencionaban así algunas de las características básicas de los tesauros como es la finalidad de facilitar la recuperación de información y el hecho de estar formado por términos que representan conceptos que se vinculan mediante distintos tipos de relaciones.

Antes de la aparición de los primeros tesauros en la década de los 60, se utilizaban vocabularios desarrollados por instituciones nacionales e internacionales, que eran plenamente conscientes de las desventajas que suponía la utilización de clasificaciones universales para la recuperación de información especializada. Estos vocabularios, por lo general, carecían de una estructura, salvo la ordenación alfabética (Gil Urdiciain, 1998a, p. 73).

Esta característica de los primeros vocabularios es un rasgo común en una de las tendencias señaladas por Lancaster (1995) en su análisis sobre la evolución de la normativa sobre tesauros. Recordemos que según este autor, en el origen y evolución de los tesauros se fusionan dos corrientes (Lancaster, 1995):

1. La que recoge la tradición de representar la información a través de palabras y expresiones, desarrollada en EE.UU. El primer tesoro se realizó a partir de un control de los términos de lo que básicamente era un sistema Uniterm. En el Departamento de Defensa de Estados Unidos se elaboró el primer tesoro en 1960, y, en 1961, el American Institute of Chemical Engineers (AIChE) publicó el suyo. La construcción de estos tesauros sirvió de base para la creación de las primeras recomendaciones sobre elaboración de tesauros. Seguidamente los departamentos de Industria y Defensa desarrollaron conjuntamente el proyecto LEX (1965-1967) cuyo resultado fue la publicación, en, de otro tesoro, el TEST. Las directrices sobre construcción de tesauros derivadas del Proyecto Lex fueron aprobadas y publicadas y se recomendó a todas las agencias gubernamentales de Estados Unidos que las adoptaran para la construcción de sus tesauros.
2. La que recoge la tradición de los sistemas de clasificación basados en concepciones lógicas que permiten agrupar los conceptos y los documentos relacionados semánticamente. Se desarrolló en Europa, principalmente en Inglaterra y Alemania. El método de la clasificación analítico-sintética por facetas creado por Ranganathan en los años 30 fue clave para el afianzamiento de esta normativa. Además, sus principios dieron pie a la creación de esquemas de clasificación especializados e influyeron en la indización alfabética de materias. Por otro lado, la concurrencia de la clasificación por facetas y del tesoro convencional quedó patente cuando se creó el tesoro por facetas en 1969 en un intento por conseguir una herramienta que agrupara lo mejor de ambas clasificaciones.

Cada corriente se desarrolló de manera independiente, pero llegado un momento ambas corrientes confluyeron y sirvieron como base para la creación de distintas normativas nacionales e internacionales. Estas normativas han establecido las recomendaciones más importantes para la creación y el mantenimiento de tesauros monolingües y multilingües. Los estándares han ido cambiando a medida que los tesauros iban evolucionando.

Por otro lado, un factor determinante en la evolución reciente de la normativa ha sido la aparición de las nuevas tecnologías y de Internet, lo que ha permitido que los tesauros se hayan adaptado a los nuevos tiempos y evolucionen desde su primitivo formato impreso y su uso para indizar y recuperar información.

En este proceso tienen un papel fundamental las instituciones que han creado estos instrumentos electrónicos y que ahora son responsables de adaptarlos a entornos electrónicos, como unas características que les permitan evolucionar hasta la web semántica. El objetivo es que

sirvan para compartir el conocimiento y que puedan utilizarse como herramientas para la búsqueda de información.

Con el formato electrónico, se quiere hacer llegar los tesauros a una comunidad de usuarios más amplia, por lo que actualmente encontramos multitud de normas que hablan de su creación, mantenimiento y estructura (Sánchez-Cuadrado et al., 2012).

La adaptación de los tesauros al entorno web ha supuesto que su manejo sea más fácil gracias a los enlaces que incluyen, que los costes de actualización sean más baratos y que se permita la participación de los usuarios en su actualización,. Además, se ofrece la posibilidad de la interoperabilidad con otras aplicaciones, por lo que se da un paso más hacia la web semántica. Pero también se observan algunas desventajas como pueden ser la falta de herramientas que conecten los términos con objetos (Mochón Bezares; Sorli Rojo, 2010, p. 644).

2. Evolución de la normativa relativa a tesauros

2.1 Primeras normas

Desde la creación de los primeros tesauros, los estándares y recomendaciones sobre ellos han ido evolucionando en el tiempo, adaptándose a las nuevas realidades que iban surgiendo.

La aparición de Internet ha marcado un antes y un después en la normativa: hemos pasado de tener normas con un objetivo principal de mejorar la consistencia de los tesauros (a través de un control terminológico y una estructura semántica) para facilitar la creación de índices consistentes, a normas donde se mantienen todos los objetivos anteriores y se añaden la facilidad de consulta, la gestión electrónica y la interoperabilidad con otras aplicaciones y herramientas. A continuación haremos un repaso de las normas más importantes y nos centraremos en la que nos ha servido como base para la realización de este trabajo, la ISO 25964-1.

Como ya hemos explicado anteriormente, y para resumir su origen, sabemos que de los primeros tesauros aparecidos en los años 60-70 en los departamentos de Defensa e Industria de EE.UU. y del Proyecto LEX, surgieron también las primeras directrices sobre su construcción. Dichas directrices sirvieron de base para la creación del tesoro de la UNESCO (1970), y sus recomendaciones permitieron, a su vez que la ISO publicara también las primeras normas en 1974.

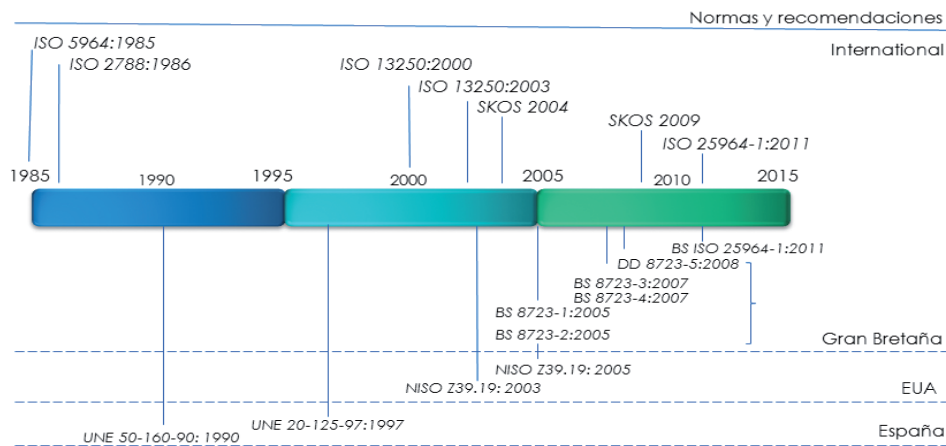


Imagen 1. Evolución de la normativa sobre tesauros (Dextre-Clarke; Lei Zeng, 2011)

-La primera norma que trató específicamente sobre los tesauros fue la ANSI Z39.19-1974 titulada *Thesaurus Structure, Construction and Use* y fue publicada por el American National Standards Institute (ANSI) en 1974. La norma se revisó en 1980, en 1993 fue analizó de nuevo y cambió de nombre y en 1998 se volvió a repasar y se confirmó. Este análisis de finales de los 90 se llevó a cabo para poner la norma al día en cuanto a los cambios que se estaban produciendo en el entorno informático y en el de la WWW.

A partir de esta revisión, la NISO organizó el “Workshop Electronic Thesauri” en noviembre de 1999 con el objetivo de crear un estándar para los tesauros electrónicos. Este taller-reunión fue considerado el punto de partida de la renovación de las normas sobre tesauros y para su adaptación al entorno digital. La reunión contó con el apoyo de muchas asociaciones, y de los acuerdos a los que se llegó surgió la cuarta versión de la norma Z39.19, que se publicó en 2005 (la NISO Z39.19-2005). Esta adaptación incluía los siguientes apartados: análisis preliminar con la importancia del control en el vocabulario, principios y estructura del tesoro, normas de control de los términos, relaciones entre los términos, visualizaciones, características de la interoperabilidad, y recomendaciones para el mantenimiento y gestión del tesoro (García Marco et al, 2007, p. 392).

Esta última versión (Z39.19-2005) está pensada para ser aplicada a tesauros monolingües e incorporaba su aplicación a entornos digitales, como demuestra el tratamiento de esquemas de metadatos como *Dublin Core* (Sánchez-Cuadrado; Colmenero-Ruiz; Moreira, 2012, p. 232). Fue completada con las directrices que la IFLA publicó en 2009 sobre la vinculación entre diferentes vocabularios controlados y la construcción de tesauros de forma no simétrica. En sus inicios, la novedad de esta norma norteamericana radicó en su nuevo enfoque dirigido tanto a recursos de información impresos y electrónicos, en su ampliación del alcance a todos los vocabularios controlados (anillos de sinónimos, taxonomías, etc.), no solo tesauros, y en el hecho de abordar el tema de la interoperabilidad en el marco de Internet (García Marco et al, 2007, p. 392)

-En 1974 se publicaba también la primera versión de la ISO 2788 *Documentation -- Guidelines for the establishment and development of monolingual thesauri* en cuya introducción se afirmaba:

“Hay una necesidad de tener métodos prácticos para representar conceptos de manera simple y clara, y de ordenarlos teniendo claras sus relaciones” (Dextre Clarke; Lei Zeng, 2011, p. 21). Esta norma está inspirada en las recomendaciones de la UNESCO. Se revisó y publicó en 1986. A nivel internacional, esta norma se ha mantenido durante 25 años hasta la aparición de la ISO 25964-1 en 2011, que la anuló. (Sánchez-Cuadrado; Colmenero-Ruiz; Moreiro, 2012, p. 231). Cada país publicó su versión: en el caso de Gran Bretaña fue la BS 5723:1979, y en el de España, la UNE 50-106-90.

-En 1985 se publicó la ISO 5964 *Documentation -- Guidelines for the establishment and development of multilingual thesauri*. Esta norma establece los procedimientos principales para el tratamiento de varias lenguas en un mismo recurso de conocimiento, identificando los principales problemas en cuanto a la equivalencia semántica. En España esta norma se tradujo en la UNE 50-125-97 y en Gran Bretaña, en la BS 6723:1985 (Sánchez-Cuadrado et al., 2012, p. 231). Esta norma ISO 5694 también fue completada con las directrices publicadas por la IFLA en 2009 y, como la ISO 2788, también ha sido anulada con la aparición de la ISO 25964-1:2011.

Durante muchos años, la ISO 2788:1974 y la ISO 5964:1985 fueron las únicas referencias normativas a nivel internacional para la creación y mantenimiento de tesauros, a pesar de que una de sus principales desventajas era la no distinción entre términos y conceptos (Pastor-Sánchez, 2013).

-Aunque, como hemos señalado, la primera norma en adaptarse a entornos electrónicos fue la americana, entre el 2005 y el 2008 se publicaron las cinco partes de la norma BS 8723 (García Marco et al, 2007, p. 393). La última parte (denominada DD8723-5) es la más interesante porque está dedicada a protocolos y formatos de intercambio para la interoperabilidad. Las cuatro primeras partes de la norma anulan la BS 5723 y la BS 6723. Ahora, toda la norma ha quedado anulada tras la publicación de la BS ISO 25964:2011, idéntica a la ISO 25964-1 de 2011.

Pero no solo han surgido normativas sobre la creación de tesauros; la interoperabilidad, de la que ya hemos hablado, necesita formatos de intercambio para que sea un proceso más fácil, de manera que se puedan traspasar a la web la información semántica de las estructuras de conocimiento. A continuación presentamos diferentes estándares y/o recomendaciones centrados exclusivamente en este aspecto:

-Entre el 2000 y el 2003 se publica la ISO 13250, cuya principal aportación es la definición de los *topic maps* (TM) en un formato basado en XML para facilitar la interoperabilidad.

-En el 2004 SKOS el W3C publica su propia recomendación, que fue actualizada y publicada en 2009. El lenguaje de formalización es RDF y se puede combinar con lenguajes formales que admiten la incorporación de otros vocabularios de metadatos. Es el formato que está más implementado actualmente, gracias a que ha sido adoptado por la iniciativa

Linked Data como lenguaje de formalización para los vocabularios de valores (Caro Castro, 2012).

2.2 Norma ISO 25964

Una vez analizadas las distintas normativas sobre tesauros, dedicaremos un apartado especial a comentar las características más importantes de la última norma publicada sobre estas herramientas, la ISO 25964, que es la que nos ha servido de base para la realización de este trabajo.

Recordamos que esta norma surge como revisión de las normas ISO 5964 e ISO 2788 y de algunas partes de la BS 8723, por lo que las sustituye. Esta revisión, que se convierte en la renovación más profunda que ha experimentado el estándar ISO de los tesauros, era necesaria porque después de 25 años esas normas ya no respondían a los nuevos modelos de indización y recuperación de la información (Sánchez-Cuadrado et al., 2012, p. 232).

La *ISO 25964 Information and documentation — Thesauri and interoperability with other vocabularies* trata, como su propio nombre indica, sobre la creación y mantenimiento de tesauros y la interoperabilidad de estas herramientas con otras aplicaciones. La norma fue preparada por el Comité Técnico ISO / TC 46: Información y Documentación; Subcomité SC 9: Identificación y descripción (ISO 25964-1, 2011, p. 7). El título general abarca dos partes:

-Parte 1: *Thesauri for information retrieval*, es decir, tesauros para la recuperación de información (que es la que se ha utilizado para el trabajo). Se publicó en 2011.

-Parte 2: *Interoperability with other vocabularies*, esto es, la interoperabilidad con otros vocabularios. Se publicó en marzo de 2013.

La primera parte ofrece recomendaciones para la construcción, desarrollo y mantenimiento de tesauros monolingües y multilingües destinados a aplicaciones de recuperación de información, e incluye recomendaciones funcionales para el mantenimiento del software del tesoro (Dextre Clarke; Lei Zeng, 2011, p. 21). Además, se aplica a los vocabularios utilizados para la obtención de información sobre todo tipo de recursos, independientemente de los medios utilizados (texto, sonido, imagen, etc). El documento también proporciona un modelo de datos, así como el formato y los protocolos recomendados para la importación y exportación de los mismos o su intercambio (ISO 25964-1, 2011, p. 1).

La segunda parte abarca la interoperabilidad entre diferentes tesauros y con otros tipos de vocabularios estructurados -como taxonomías, listas de autoridades, ontologías, esquemas de clasificación, anillos de sinónimos, etc. Como se aprecia, el intercambio eficiente de datos se está convirtiendo en un componente esencial de la gestión de un tesoro y de su explotación, por lo

que esta norma incluye recomendaciones para la realización de esta función (ISO 25964-1, 2011, p. 7).

Los principales cambios que presenta esta norma con respecto a las anteriores es que fusiona la normativa para tesauros mono y multilingües, propone un formato de intercambio basado en un esquema XML, introduce un cambio en el formato (de impreso a digital), aumenta el uso de facetas para la organización del conocimiento y consigue adaptarse a los nuevos perfiles de los usuarios. Es mucho más rigurosa que sus antecesoras y distingue claramente los términos y los conceptos (Sánchez-Cuadrado et al., 2012, p. 232).

A la vista de la evolución normativa, observamos que el uso de los tesauros ha ido variando con el paso del tiempo: anteriormente se diseñaron para profesionales de la información y la documentación que eran expertos en indización y búsqueda, pero hoy en día el diseño se dirige a vocabularios que permitan la inferencia con las máquinas. Actualmente los tesauros son, principalmente, herramientas electrónicas que se construyen y mantienen con el apoyo de un software y que, para ser utilizados, necesitan integrarse con otro software, como los motores de búsqueda y los sistemas de gestión de contenidos (ISO 25964-1, 2011, p. 8).

Sin embargo, aunque estas herramientas siguen teniendo un claro uso en la indización y recuperación de documentos, el reto más importante al que nos enfrentamos los profesionales es conseguir que estos vocabularios sean útiles a los usuarios finales en la búsqueda de información, una función que debería haberse desarrollado con la aparición de las nuevas tecnologías de información pero que no ha terminado de despegar. A pesar de que tanto la bibliografía profesional como los organismos internacionales ya han comenzado a darse cuenta de esta funcionalidad, todavía queda mucho camino por recorrer para que se haga realidad (Caro Castro, 2012).

Esta norma dedica un apartado muy importante a cómo deben presentarse las relaciones entre conceptos y entre términos y conceptos en los tesauros, es decir, las visualizaciones. Este es uno de los aspectos que ha incorporado cambios sustanciales respecto a la norma española UNE 50-106-90, que es la que comentaremos a continuación. A la evolución del apartado de la visualización le dedicaremos un capítulo aparte en el trabajo.

2.3 Norma UNE 50-106-90 y UNE 50-125-97

Aunque la ISO 25964-1 ha sido la norma base para la realización del trabajo, consideramos necesario señalar unos pequeños apuntes sobre las normas españolas relativas a los tesauros monolingües y multilingües.

Como se lee en la evolución normativa, AENOR ha ido siempre con mucho retraso temporal a la hora de adaptar las normativas internacionales al caso español. Recordemos que la UNE 50-106 se publicó en 1990 como una adaptación de la ISO 2788, publicada en 1974; o que la UNE 50-125

se publicó en 1997, doce años después de la norma en la que se basa, la ISO 5964, que se publicó en 1985. Sin embargo, desde hace unos años esta entidad se ha sumado con rapidez a los esfuerzos internacionales para revisar y ampliar el alcance de las normas sobre tesauros, y así facilitar su adaptación al entorno español. Así, el 23 de junio de 2006 se constituyó oficialmente el grupo de Vocabularios Controlados para la Recuperación de Información del Comité 50, cuyo objetivo principal es la reunificación de las normas relativas a tesauros monolingües y multilingües en una sola (García Marco, et al, 2007, p. 393-394)

-Norma UNE 50-106-90: Esta norma trata aspectos relacionados con la selección de términos (ya que incluye recomendaciones para el control del vocabulario), pero se ocupa especialmente de los medios para establecer y presentar los distintos tipos de relaciones entre términos de indización. Las recomendaciones no deben considerarse como instrucciones de obligado cumplimiento, ya que en muchos casos se describen procedimientos opcionales que variarán según el centro donde se apliquen. En cambio, sí es cierto que las técnicas descritas están basadas en principios generales aplicables a cualquier campo temático. La norma contiene recomendaciones relativas a la construcción y mantenimiento tesauros monolingües, sin referencias a los requisitos propios de los tesauros multilingües, aunque han sido aceptadas como bases para la creación de la norma de estos últimos (UNE 50-106, 1990).

-Norma UNE 50-125-97: La base que se utilizó para la creación de esta norma fue la UNE 50-106-90 para tesauros monolingües, por lo que las directrices son una extensión de esa norma. Los principios son prácticamente los mismos, y la principal novedad que incluye es la existencia de otras lenguas a parte del español. De esta manera, cada término preferente tendrá su equivalencia interlingüística en los idiomas del tesoro (UNE 50-125, 1997).

3. Visualización y navegación de los tesauros e integración con las aplicaciones

Como destacan Dextre Clarke y Lei Zeng (2012), la nueva norma internacional contribuye a aclarar la confusión que se pudiera haber generado en estándares anteriores entorno a la diferenciación entre término y concepto. La noción de concepto frente a la de término es el eje que vertebra el modelo estructural de los tesauros: son estos los que se organizan gracias a relaciones jerárquicas y/o asociativas; y los que se agrupan, identificados (o no) por una etiqueta de nodo. Para las personas, los conceptos se denominan mediante etiquetas preferentes o alternativas (términos) en una o más lenguas; mientras que para las máquinas esta denominación es un URI.

Si esta es una de las aportaciones importantes de la norma, otro eje fundamental es la gestión automatizada del tesoro. De todos los aspectos que puede abarcar esta gestión – creación, mantenimiento, actualización, interoperabilidad, uso – nos parecen fundamentales aquellos relacionados con la potencial interacción del usuario: la visualización, la búsqueda de términos dentro del propio tesoro y la conexión con bases de datos para recuperar información

(integración con las aplicaciones). En 2007, dos años después de la publicación de las dos primeras normas adaptadas, Williamson reconocía dos tipos de visualizaciones en los tesauros electrónicos:

- **Visualización estática:** es idéntica a la del formato impreso, aunque el tesoro tenga acceso electrónico. Los contenidos se pueden explorar aunque hay pocas facilidades para moverse por el tesoro de forma dinámica. A nivel práctico son poco útiles. Este tipo de presentación fue la utilizada en los primeros intentos de pasar de un instrumento impreso a uno electrónico.
- **Visualización dinámica:** en este modelo hay más facilidades para la búsqueda y la navegación. Se ofrecen a los usuarios distintas formas de acceso a los conceptos: a partir de una lista de categorías, de búsquedas booleanas o de hipervínculos para moverse entre las diferentes visualizaciones que ofrecen los tesauros. Algunos tesauros proceden de formatos impresos y otros se han creado exclusivamente para usarse en el entorno digital.

Dos años después de haberse publicado la primera parte de la ISO 25964, no parece cuestionable que los vocabularios estructurados se gestionan ya de forma automatizada. Tampoco que ha acelerado su integración en la web semántica, editándose en formatos – como SKOS – que posibilitan la interoperabilidad y su uso en la descripción de documentos con metadatos (Pastor-Sánchez; Martínez Méndez; Rodríguez Muñoz, 2009b). Sin embargo, ¿en qué medida se ha mejorado su presentación?, ¿se han incorporado herramientas para facilitar la interacción con el usuario? Estos son los aspectos en los que se centra este trabajo y a los que dedicaremos los siguientes apartados.

3.1 Visualización

Las diferentes presentaciones o visualizaciones de un tesoro responden a la necesidad de facilitar su utilización para la indización y para la búsqueda. También al objetivo de mostrar los términos (descriptores o no), aclarar su significado y evidenciar las relaciones conceptuales. A continuación expondremos las características de los distintos tipos de visualización comparando, cuando sea posible, las recomendaciones de la norma ISO 25964-1 con las de su antecesora la ISO 2788, concretamente de su traducción al español: la UNE 50-106-90¹.

Por su parte, en la ISO 25964-1 se considera que, cuando se presenta para un uso humano, ya sea en una pantalla o en forma impresa, el tesoro se puede mostrar de las siguientes maneras:

Un **registro simple** (apartado 12.2.2) que es la forma más elemental de visualización, mostrando el término preferente o no preferente y, opcionalmente, alguna o todas las relaciones

¹ Las referencias a la ISO 5964:1985 (UNE 50-125:1997) solo se harán en aquellos casos en que sea necesaria una referencia explícita al multilingüismo.

vinculadas a ese concepto: código o notación, notas (de alcance, de historia o de definición), referencias hacia términos equivalentes, referencias al término cabecera, a términos genéricos o específicos, a otros términos relacionados y al grupo en el que se incluye el concepto (microtesauro o campo temático).

Para la presentación impresa, la UNE 50-106-90 prevé que toda la información semántica de un término se incluya en la *presentación alfabética estructurada*, que incluye los términos preferentes y no preferentes en una secuencia alfabética única.

mercado financiero
 MT 2421 libre circulación de capitales
 UF mercado de valores
 UF mercado financiero internacional
 NT1 bolsa de valores
 NT2 cambio de los valores bursátiles
 NT2 cotización en bolsa
 NT3 valor bursátil
 NT1 emisión de valores
 NT1 estadística financiera
 NT1 instrumento financiero
 NT1 intervención financiera
 NT1 mercado a término
 NT1 mercado al contado
 NT1 operación de bolsa
 NT1 profesión financiera
 NT1 situación financiera
 NT1 transacción financiera
 NT1 valor mobiliario
 NT2 valor de renta fija
 NT2 valor de renta variable
 RT institución financiera
 RT mercado monetario
 RT oferta pública de compra
mercado financiero internacional
 USE **mercado financiero**
mercado inmobiliario

mercado financiero

SN
 Lugar de emisión e intercambio de valores mobiliarios,
 principalmente acciones y obligaciones.

UF actividad financiera
 mercado de valores
 mercado financiero internacional

24 ASUNTOS FINANCIEROS

MT 2421 libre circulación de capitales
 NT1 Bolsa de valores
 NT2 cotización bursátil
 NT3 valor bursátil
 NT2 cotización de valores
 NT1 emisión de valores
 NT1 estadística financiera
 NT1 instrumento financiero
 NT1 intervención financiera
 NT1 mercado al contado
 NT1 mercado a plazo
 NT1 operación bursátil
 NT1 profesión financiera
 NT1 situación financiera
 NT1 transacción financiera
 NT1 valor mobiliario
 NT2 valor de renta fija
 NT2 valor de renta variable
 RT institución financiera [2416]
 mercado monetario [2411]
 oferta pública de adquisición [4006]

EQUIVALENTES LINGÜÍSTICOS

BG финансов пазар
 ES **mercado financiero**
 CS finanční trh
 DA finansielt marked
 DE Finanzmarkt
 ET finantsturg
 EL πιστωτική αγορά
 EN financial market
 FR marché financier
 HR financijsko tržište
 IT mercato finanziario
 LV finanšu tirgus
 LT finansų rinka
 HU pénzügyi piac
 MT financial market (under translation)
 NL geldmarkt
 PL rynek finansowy
 PT mercado financeiro
 RO piață financiară
 SK finančný trh
 SL finančni trg
 FI rahoitusmarkkinat
 SV finansmarknad
 SR финансијско тржиште

Imagen 2. Tesauro EUROVOC (1990). Lista alfabética estructurada (fragmento)

Imagen 3. Tesauro EUROVOC (en línea). Registro simple

Respecto a la *presentación sistemática*, la norma española (equivalente a la ISO 2788) indica que la organización primaria del tesauro puede ajustarse a dos enfoques diferentes: temas y facetas. Su visualización impresa debe incluir, como mínimo, los descriptores organizados jerárquicamente. Se pueden añadir términos equivalentes y relacionados, cuando no haya presentación alfabética estructurada e indicadores clasificatorios (indicadores de faceta) para mostrar el criterio de clasificación de una categoría.

La ISO 25964-1 diferencia entre una **visualización jerárquica** y una **visualización clasificada** en función de que, en un primer nivel, el tesauro esté organizado por facetas o por temas. Cada uno de los enfoques tiene ventajas y desventajas, en función del contexto de la aplicación. Sin embargo, es importante que esté disponible al menos una de las formas de visualización.

La visualización jerárquica se considera más apropiada para los tesauros que se limitan a un campo temático y que están organizados por facetas. En este caso las relaciones entre términos específicos y términos genéricos responden en sentido estricto a los criterios normativos, por lo que ayudan en la generalización o especificación del concepto indizado o buscado. En esta forma de visualización normalmente no se muestran otras relaciones (por ejemplo, las de equivalencia y las asociativas), ya que harían que la estructura fuera difícil de seguir.

La visualización clasificada sería la indicada para aquellos vocabularios estructurados en los que el enfoque de la macroestructura del sistema es por disciplinas o campos temáticos. Este punto de vista tiene la ventaja de organizar los conceptos en grupos que generalmente corresponden a los modos de pensar de los usuarios. El inconveniente es que algunos conceptos se utilizan habitualmente en varios campos diferentes, por ejemplo, si un tesoro incluye la silvicultura y la agronomía, habrá muchos conceptos, como la plantación, que se produzcan en ambos.

Con la agrupación de conceptos por campo temáticos, se producen visualizaciones que se parecen a las jerarquías pero en las que los términos subordinados pueden no seguir las reglas para las relaciones jerárquicas. En este caso plantea la posibilidad de que se utilice una organización en facetas dentro de los campos temáticos, indicándose mediante etiquetas de nodo (indicadores de faceta) los diferentes criterios clasificatorios, y dónde las relaciones jerárquicas no se ajustan exactamente a la norma.

En cualquiera de los dos casos se establece que la visualización sistemática de un idioma se construye colocando una sangría a cada rango de términos específicos bajo su término genérico. Es posible añadir a este tipo de visualización todos los datos que se describen para el registro simple, aunque la incorporación de datos adicionales hace que la presentación sea más extensa.

5206 política del medio ambiente

gestión del agua

RT contaminación del agua [5216]
RT protección de las aguas [5206]
RT recurso hidráulico [5211]

NT1 análisis del agua

NT1 consumo de agua

RT agua potable [5211]
RT consumo alimentario [2026]

NT1 necesidades de agua

RT necesidades fundamentales [1611]

NT1 tratamiento del agua

RT agua residual [5216]
RT lodos de depuración [5216]

NT1 utilización del agua

RT agua [5211]

gestión de residuos

RT exportación de residuos [5216]
RT lucha contra el despilfarro [5206]
RT residuo [5216]

NT1 almacenamiento de residuos

NT2 almacenamiento subterráneo de residuos

NT1 eliminación de residuos

RT biodegradabilidad [5206]

NT2 incineración de residuos

NT2 inmersión de residuos

RT contaminación marina [5216]

NT1 reciclaje de residuos

RT depósito por producto contaminante [5206]
RT industria de pasta y papel [6836]
RT residuo de la madera [6836]

9 entradas correspondientes a : Derechos humanos

Para obtener las noticias indizadas con un término dado en unesdoc/unesbib, haga click en el [número]

Término : **Derecho humanitario** [222]

MT 6.10 Derechos humanos
UP Derecho de la población civil
UP Derecho de las víctimas de guerra
UP Derecho de los conflictos armados
UP Derecho de los prisioneros de guerra
UP Derecho internacional humanitario
TE Crimen de guerra [82]
UP Criminal de guerra
....TE2 Campo de concentración [21]
UP Campo de exterminio
....TE2 Genocidio [124]
TE Territorio ocupado [212]

Término : **Derechos civiles** [613]

MT 6.10 Derechos humanos
UP Derecho a la dignidad
UP Derechos políticos
UP Libertad
UP Libertades civiles
UP Libertades políticas
TE Derecho a buscar asilo [67]
TE Derecho a la justicia [57]
UP Derecho a juicio
UP Derecho a un proceso imparcial
UP Derecho a un proceso justo
UP Garantías jurídicas
TE Derecho a la privacidad [211]
UP Confidencialidad
UP Derecho a la imagen
UP Derecho a la intimidad
UP Protección de la vida privada
UP Respeto de la vida privada
TE Derecho a la propiedad [35]
UP Legislación sobre la propiedad
UP Legislación sobre los bienes
TE Derecho a la vida [47]

Imagen 4. Tesoro EUROVOC (en línea). Visualización sistemática

Imagen 5. Tesoro de la UNESCO (en línea) Visualización sistemática

Otro tipo de presentación es un **listado alfabético**, que permite el acceso a los conceptos desde las palabras en las que están inicialmente expresados por el usuario. En un tesauro impreso actúa como un índice, reenviando a la presentación sistemática mediante un código (UNE 50-106-90) y en la pantalla de ordenador puede complementar una función de búsqueda directa. En una visualización alfabética todos los términos, ya sean preferentes o no preferentes, están organizados como una secuencia alfabética única, distinguiéndose tipográficamente de modo que la diferenciación entre ellos sea evidente. Este listado puede limitarse a una sencilla lista de términos – permutada o no – o incluir información complementaria: términos genéricos o específicos más inmediatos o el registro completo de cada término.

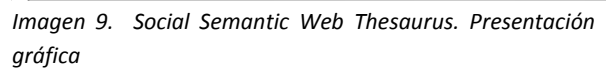


Imagen 6. Tesauro EUROVOC (1990). Índice alfabético



Imagen 7. Tesauro EUROVOC (en línea). Visualización alfabética

La UNE 50-106-90 considera que en la **presentación gráfica** los términos de indización y sus relaciones se disponen en forma de una figura bidimensional que permite examinar una porción de términos interrelacionados. Se identifican dos tipos principales de presentación gráfica: estructura arborescente y diagrama de flechas. En la parte gráfica solo se recogen los términos preferentes y se remite a la sección alfabética estructurada mediante algún sistema de reenvíos. Las indicaciones de la ISO 25964-1 respecto a este tipo de visualización son muy genéricas: se indica que puede dar una visión general de un campo que resulte útil, pero es compleja de crear y actualizar, y resultará confusa si contiene demasiada información adicional.



3.2 Navegación y búsqueda

La visualización en línea puede hacer que el proceso de búsqueda sea mucho más fácil (Pastor-Sánchez; Saorín Pérez, 1996). Los enlaces explícitos entre conceptos no serán necesarios cuando el tesauro se muestre en una pantalla de ordenador porque los hipervínculos generados automáticamente deberían:

- Cambiar de un tipo de presentación a otra a partir del descriptor seleccionado. Preferiblemente, debería ser posible verlas simultáneamente.
- Moverse desde el registro de un concepto hasta cualquiera de sus términos más genéricos, más específicos o que están relacionados.
- Expandir y contraer los niveles de las visualizaciones jerárquicas o clasificadas, mostrando tantos detalles como los usuarios necesiten sólo cuando lo quieran ver.
- Elegir o cambiar el idioma de la pantalla de un tesauro multilingüe, o limitar la búsqueda y mostrar un solo idioma especificado.

Como complemento a las opciones de navegación, los tesauros tendrían que facilitar el acceso a los conceptos mediante algún sistema que permitiera realizar búsquedas analíticas. Según la ISO 25964-1, las aplicaciones de búsqueda deberían incluir la posibilidad de:

- Buscar por términos que contienen cualquier palabra, una combinación de palabras o una cadena de caracteres.
- Buscar conceptos utilizando cualquier notación y / o identificadores
- Mostrar los términos equivalentes (preferentes o no preferentes) para un término dado, incluyendo las equivalencias interlingüísticas si el tesauro es multilingüe.

3.3 Integración con las aplicaciones

La función principal de un tesauro es apoyar la recuperación de información. El tesauro suele incidir en dos etapas principales en la cadena de recuperación: en el momento de la indización y en el momento de la búsqueda. Ahora el escenario se ha vuelto más complejo con la difusión de las redes electrónicas y con el uso casi universal de los ordenadores para gestionar información. Esto hace que se presenten oportunidades de compartir datos e información en una escala nunca antes imaginable: desde un solo ordenador conectado a Internet uno puede buscar al mismo tiempo diferentes colecciones de libros, artículos, fotografías, datos u objetos de museo que se encuentran en equipos remotos de todo el mundo.

Para permitir la explotación de los tesauros en este tipo de escenarios, los sistemas implicados necesitan interoperabilidad, es decir la capacidad de múltiples sistemas o componentes para intercambiar y utilizar información. Pero además, para ser realmente funcionales, deben estar integrados en el sistema de búsqueda que se utilice en una base de datos o en un repositorio documental.

Respecto a la interoperabilidad necesaria para los tesauros, para hacer que un tesauro esté disponible para los usuarios de una red como la de Internet, debería ser publicado en esa red, utilizando un formato estándar. Además, debería ser posible exportar e importar un tesauro, en parte (registros nuevos o modificados) o en su totalidad, desde / hasta la aplicación en la que se mantiene.

Si un tesoro se va a utilizar para la recuperación de la información, debe ser capaz de integrarse con los sistemas en los que la indización y búsqueda se llevan a cabo. Para un tesoro impreso, la integración se logra mediante el proceso de una persona que mira a través de las páginas, seleccionando los términos y volviendo a meterlos en el sistema de recuperación de la información, con la sintaxis correcta. Para un tesoro electrónico, estas funciones deben ser apoyadas automáticamente de manera que, mientras que navega por el tesoro, el usuario pueda ser capaz de pasar directamente a la fase de búsqueda. Se deberían facilitar las siguientes opciones:

- Posibilidad de seleccionar uno o más términos para utilizarlos en una búsqueda, sin necesidad de volverlos a escribir.
- Cuando un usuario introduce un término que tiene varios significados en el tesoro, el motor de búsqueda puede responder mediante la recuperación de resultados de todos estos términos, agrupándolos por los conceptos que representan. Por ejemplo, si un tesoro tiene dos términos "Venus (Divinidad griega)" y "Venus (Planeta)", y el usuario busca simplemente por "Venus", los resultados de la búsqueda deberían presentarse en dos grupos, correspondientes a cada uno de los conceptos del tesoro, respectivamente.
- La selección de un término no preferente debería conducir a la sustitución por el término preferente en el estado de búsqueda (asegurando que el usuario ha sido informado de lo que se ha hecho). En el caso de que el usuario introduzca un término no preferente en un enunciado de búsqueda sin antes navegar por el tesoro, se le ofrecerá la opción de convertirlo en el término preferente correspondiente.
- Las relaciones de equivalencia en un tesoro también pueden ser explotadas para buscar una base de datos que no haya sido indizada con un tesoro. Por ejemplo, si un tesoro incluye "motos" y "motocicletas", entre los términos preferentes y no preferentes para el mismo concepto, se puede ampliar una búsqueda por cualquiera de estos términos para incluir los otros.
- Mostrar el número de documentos a los que se ha asignado cada término preferente. Durante el proceso de refinamiento de búsqueda, el número de publicaciones en cualquiera de las pantallas del tesoro debería ajustarse para mostrar los números que se aplican al subconjunto de la colección seleccionada.
- Posibilidad de construir una estrategia de búsqueda facilitando algún tipo de ayuda con el uso de la sintaxis (por ejemplo, seleccionando operadores booleanos).
- Posibilidad de seleccionar subárboles desde el tesoro, es decir, la búsqueda de un término dado y de todos sus términos específicos combinados con el operador booleano OR, a veces llamada función de "explode".

III. ESTUDIO

1. Objetivos

El objetivo principal de este trabajo de investigación es evaluar la interfaz de consulta, las opciones de navegación y búsqueda y la integración de tesauros con otras aplicaciones, tomando como base las características que, sobre estos aspectos, se recogen en la norma ISO 25964-1, cuyo título es *Information and documentation — Thesauri and interoperability with other vocabularies*

Como objetivos específicos podemos señalar:

- Identificar los distintos tipos de visualizaciones y los elementos que deben incluir los tesauros electrónicos según esta norma.
- Comprobar si los tesauros en formato electrónico cumplen o no la normativa internacional en cuanto a presentación y diseño.
- Comprobar si los tesauros en formato electrónico cumplen o no la normativa internacional en cuanto a búsqueda y navegación.
- Comprobar si los tesauros en formato electrónico cumplen o no la normativa internacional en cuanto a integración con otras aplicaciones desde la página principal del tesoro.
- Averiguar cuáles son los tesauros más y menos eficientes en función de los aspectos analizados

2. Metodología

El primer paso fue la lectura y revisión bibliográfica de tres temas diferentes: el primero, relativo a aquellos artículos más afines al objeto principal de este trabajo, es decir, artículos relacionados principalmente con los estándares, visualizaciones e integración de los tesauros; el segundo, relacionado con la evolución de la normativa sobre tesauros; y el tercero, sobre la integración de tesauros con otras aplicaciones. La búsqueda de artículos tuvo lugar en bases de datos especializadas como Dialnet o LISA y en páginas web oficiales como la de ISO para localizar las normativas.

La revisión más importante fue la de la norma que hemos utilizado como base para este trabajo, la ISO 25964-1. Analizamos todos los apartados que contenía e identificamos aquellos que más relación tenían con la investigación. Una vez finalizada la revisión y contextualizados los conceptos básicos, se procedió a la traducción de tres puntos de la ISO 25964-1, que consideramos clave para la posterior realización del cuestionario.

- 12 Presentación y diseño
- 13.5.1 La integración con un sistema electrónico
- 16 Integración de los tesauros con las aplicaciones que los utilizan

Tras la traducción de la norma, seleccionamos la muestra de tesauros que íbamos a evaluar. Tras barajar varios criterios, decidimos que escogeríamos aquellos que aparecieran en la wiki del *W3C Semantic Web: SKOS Datasets* por considerarla una web de referencia dentro del contexto general de mi trabajo, por la variedad y heterogeneidad temática que presentaba y por contener tesauros que estaban disponibles en libre acceso o en la web y en formato SKOS, lo que garantiza su interoperabilidad. Del listado de 39 conjuntos de datos que aparecen hemos escogido aquellos que contienen la palabra “tesauro” dentro del título o dentro del resumen para no tener ninguna duda sobre lo que nos íbamos a encontrar. El listado definitivo de tesauros evaluados es el siguiente²:

- STW Thesaurus for Economics (STW): tesauro para la economía y los negocios en inglés y alemán, que incluye una clasificación de categorías temáticas. Lo mantiene el ZBW - Leibniz Centre for Economics. Fue publicado como páginas RDFa y RDF / XML de datos, bajo licencia Creative Commons (by-nc-sa). También están disponibles un servicio web STW y SPARQL (en estado beta). STW se asigna a DBpedia, SWD, TheSoz y Agrovoc.
- Thesaurus for the Social Sciences: está disponible en alemán, francés e inglés y lo mantiene el GESIS - Leibniz Institute for Social Sciences. Contiene cerca de 12.000 entradas, de las cuales más de 8.000 son descriptores y alrededor de 4.000, no descriptores. Están incluidos los temas de todas las disciplinas de las ciencias sociales. Este conjunto de datos utiliza SKOS-XL, así como extensiones propias definidas con el fin de representar las relaciones de equivalencia entre los términos. Contiene enlaces a DBpedia, STW y AGROVOC. Se puede encontrar información adicional en la página Data Hub. Se puede descargar como RDF/XML (después de rellenar el formulario) y también está disponible a través de SPARQL. Se ha establecido una interfaz de datos vinculados en <http://lod.gesis.org/thesoz/>. El tesauro está bajo licencia Creative Commons.
- GEMET General Multilingual Environmental Thesaurus (GEMET): el Observatorio en Red de Información Europea de Medio Ambiente y (EIONET) acoge GEMET, un servicio de datos vinculados en 28 idiomas que utiliza SKOS. Disponible en <http://eionet.europa.eu/gemet>.
- AGROVOC Agricultural Thesaurus (AGROVOC): la FAO proporciona una versión en SKOS de AGROVOC través del servidor Concept disponible en <http://aims.fao.org/>. AGROVOC se ha integrado con CAT, NAL, SWD, GEMET. Algunos de estos enlaces se reflejarán pronto en el servidor.

² Disponible en: <http://www.w3.org/2001/sw/wiki/SKOS/Datasets> [Consulta: 8 de marzo]

Hubo un problema con el Cadastre and Land Administration Thesaurus ya que en el momento de evaluarlo el enlace nos redirigía a una página web que no existía y no conseguimos encontrarlo de ninguna manera, por lo que se eliminó. Tiempo después descubrimos que había estado en mantenimiento durante un largo período de tiempo y cuando se volvió a habilitar ya era demasiado tarde para incluirlo en la investigación.

- UKAT UK Archival Thesaurus (UKAT): descripción en RDF del tesauro publicado por el proyecto UKAT utilizando la versión 2005 de SKOS.
- Government of Canada Core Subject Thesaurus (GCC): tesauro bilingüe que agrupa la terminología que representa todos los ámbitos tratados en los recursos de información del Gobierno de Canadá. Disponible en: <http://en.thesaurus.gc.ca>
- Reegle Clean Energy and Climate Change Thesaurus (Reegle): contiene información sobre las energías renovables, la eficiencia energética y el cambio climático para la reutilización pública. Este tesauro es utilizado por el Linked Clean Energy Dataset.
- Social Semantic Web Thesaurus (SSW): incluye personas, organizaciones, aplicaciones y tecnologías relevantes para el área de la Web Semántica Social.
- UNESCO Thesaurus (UNESCO): lista controlada y estructurada de términos para el análisis temático y la búsqueda de documentos y publicaciones en los campos de la educación, la cultura, las ciencias naturales, las ciencias sociales y humanas y la comunicación e información.
- Eurovoc (EUROVOC): tesauro multilingüe de la Unión Europea. Los usuarios registrados pueden descargarlo en SKOS. La ontología utilizada se extiende a SKOS y SKOS-XL para expresar la estructura de dominios y microtesauros.
- Tesauro de Biblioteconomía y Documentación del IEDCYT: tesauro perteneciente al CSIC que contiene los términos preferentes y no preferentes de ese campo de conocimiento. Está disponible en español.

Para completar la lista, decidimos incluir un tesauro en español. Escogimos el tesauro de Biblioteconomía y Documentación del IEDCYT (Instituto de Estudios Documentales sobre Ciencia y Tecnología). Optamos por este tesauro por dos razones: porque el responsable de su creación y mantenimiento es CSIC, que es la mayor institución pública dedicada a la investigación en España, y por la afinidad temática con nuestra área de trabajo.

Una vez elegidos los tesauros, a continuación elegimos la técnica que emplearíamos para evaluarlos. Nos decantamos por el cuestionario por ser una herramienta eficaz y sencilla para nuestro objetivo.

Ya escogidos los tesauros y decidido la técnica de recogida de datos, seguidamente determinamos las variables que evaluaríamos. Este paso fue sencillo, pues se decidió que se tomarían de los apartados traducidos de la norma ISO 25964-1, aunque es necesario añadir que la organización de las preguntas en el cuestionario no sigue al pie de la letra la estructura de la normativa, sino que las preguntas se han reagrupado en bloques (visualización, búsqueda y navegación, e integración con las aplicaciones) siguiendo un criterio lógico para facilitar la consulta de los tesauros (ver Anexo I: Cuestionario). Una vez elegidas las variables, creamos un cuestionario de 42 preguntas pre-codificadas, cerradas y en la mayoría de los casos, dicotómicas y de respuesta indirecta

Finalmente se pasó el cuestionario a los 11 tesauros, buscando siempre 10 términos en cada uno de ellos para estar seguros de las respuestas. Estas se iban tabulando con un no o un sí (convertidos en 0 y 1, respectivamente) en tablas de Excel. A partir de esa información se obtuvieron datos descriptivos para todas las variables y se realizaron gráficos explicativos.

3. Resultados y análisis

A continuación presentamos los resultados del cuestionario que hemos pasado a los once tesauros. Los resultados pueden aparecer en forma de tabla, en forma de gráfico, combinando ambas técnicas o mediante una imagen, según la información que se quiera presentar. El objetivo es siempre que los lectores puedan interpretar fácilmente los datos, prestando especial atención a la presentación gráfica de los resultados relacionados con las visualizaciones, búsqueda, navegación o integración con las aplicaciones.

Como se observará, la mayor parte de los gráficos comparan resultados totales con las distintas variables, pues consideramos que son más importantes los resultados generales que los individuales de cada uno de los tesauros.

3.1 Visualización

3.1.1 Tipos de visualización

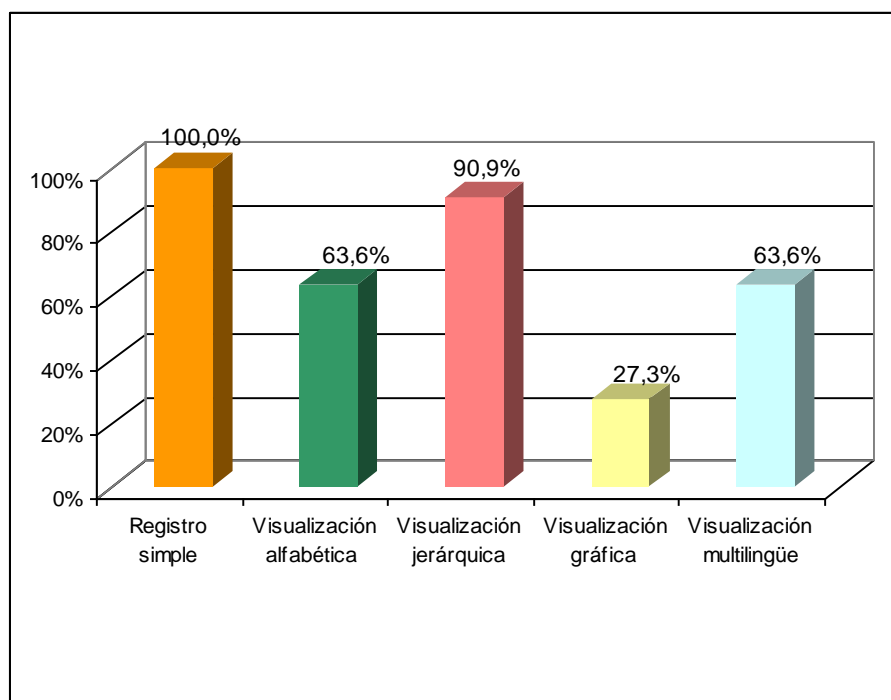


Gráfico 1. Tipos de visualización

Empezamos por un gráfico básico para nuestro trabajo, en el que se muestra en qué medida los tesauros analizados tienen las visualizaciones previstas por la norma ISO 25964-1. La misma información se presenta a continuación, con cifras absolutas, en forma de tabla.

Registro simple	11
Visualización alfabética	7
Visualización jerárquica	10
Visualización gráfica	3
Visualización multilingüe	7

Tabla 1. Tipos de visualización

Como se aprecia, lo único que cumplen todas las herramientas es tener un registro simple de cada uno de los conceptos que integran el tesoro. Consideramos que es una visualización básica porque permite que los usuarios conozcan todos los términos con los que se puede representar un concepto, las relaciones que existen entre los conceptos y las aclaraciones necesarias para entender su significado.

Le sigue la visualización jerárquica (90,9%). Este resultado puede resultar sorprendente pues en la organización conceptual de cualquier tesoro subyace una estructura sistemática. Sin embargo, se explica porque una de las herramientas analizadas – Reegle – es un glosario, a pesar de incluir en término tesoro en su nombre oficial. A favor de este resultado diremos que, como se aprecia en la tabla, sólo un tesoro no cuenta con esta visualización.

A continuación observamos que coinciden el número de tesauros que tienen presentación alfabética y multilingüe. Respecto a la primera, puede parecer extraño que no aparezca en todos los tesauros, cuando constituye una forma de acceso imprescindible para recuperar la información de los registros simples. En realidad, todos los tesauros facilitan alguna forma de acceso alfabético a la información. Solo el 63,9% lo hacen mediante un listado alfabético desplegable que el usuario puede explorar hasta encontrar el término que busca. El resto utiliza alguna herramienta de “búsqueda en el tesoro”, como se explicará más detalladamente en el apartado “Visualización alfabética”.

En cuanto a la visualización multilingüe (63,6%), presente solo en siete tesauros, creemos que en un entorno como Internet, donde conviven todas las lenguas del mundo, es imprescindible que se establezcan equivalencias interlingüísticas del contenido existente en él. Por eso lo consideramos una presentación básica para estar al día de las nuevas demandas de y necesidades de los usuarios y todos los tesauros deberían contar con ella. Si los tesauros solo eligen un idioma, perderán uso y visibilidad.

A mucha distancia de las otras visualizaciones se encuentra la presentación gráfica, solo presente en tres de los once tesauros (27,3%). Este es un resultado que sí esperábamos, pero que no entendemos. Según lo que hemos podido comprobar y lo que dice la norma UNE 50-106-90 (la

ISO 25964-1 no da una explicación sobre esta visualización), este tipo de presentación complementa el resto de las visualizaciones por lo que sería de gran ayuda para los usuarios contar con ella para aclarar posibles dificultades encontradas en las otras visualizaciones.

En los siguientes apartados se valora la adecuación de la información en cada una de estas visualizaciones a los requisitos establecidos en la ISO 25964-1.

3.1.2 Visualización del registro simple

Código del concepto	1
Notas	9
Términos no preferentes	9
Término cabecera	6
Términos generales	11
Términos específicos	11
Términos relacionados	11
Microtesauro	3

Tabla 2. Registro simple

Antes de comenzar con el análisis, recordemos que los 11 tesauros incluyen la visualización del registro simple porque su objetivo es contextualizar un concepto en un campo semántico mostrando sus relaciones (pertenencia, jerárquicas, asociativas) con otros conceptos. También muestra las notas y aclaraciones necesarias para entender su significado, así como las posibles “etiquetas” para denominar ese concepto en uno o varios idiomas (relaciones de equivalencia).

Aunque en el gráfico se enumeran 8 elementos, en realidad tiene 10. Aparecen menos porque se han agrupado tres bajo el ítem “notas”: la definición del concepto, la nota de historia y la nota de alcance. La razón es que en algunos tesauros no se diferencia el tipo de nota o se utiliza la nota de alcance para incluir una definición. En los únicos tesauros en los que se distinguían bien eran en AGROVOC y GEMET, que diferenciaban perfectamente la nota de alcance de la definición del término preferente; y en Reegle, donde, al ser un glosario, sólo había definiciones de los términos. Para que estas divergencias no distorsionaran los resultados hemos considerado mejor incluir todas las aclaraciones en el mismo epígrafe.

Un dato curioso sobre este ítem es que en el Tesauro de Biblioteconomía y Documentación del IEDCYT y STW Economics no se han encontrado ningún tipo de notas.

Respecto a la relación de equivalencia, no parece lógico que los términos no preferentes solo se encuentren en nueve tesauros (81,8%) (Social Semantic Web y GEMET no los incluyen). Estos términos son la base de las relaciones de equivalencia y son imprescindibles para que el usuario pueda utilizar cualquier expresión (o una gran variedad de expresiones) para referirse a un concepto.

Los términos generales, específicos y relacionados aparecen en todos los tesauros. Es un hecho lógico ya que son básicos para entender las relaciones semánticas de los conceptos.

Los términos cabecera, los microtesauros y los códigos del término sirven para incluir el concepto en una jerarquía. Por tanto, en cualquier visualización, refuerzan la información sobre la familia o jerarquía en la que aparece el concepto. Es imprescindible que en el registro simple se de información sobre, al menos, un nivel de jerarquía. Otras informaciones pueden mejorar la visión de conjunto pero, en sentido estricto, no serían imprescindibles. Los primeros aparecen en el 54,6% de los tesauros consultados, el microtesauro, en el 27,3% (EUROVOC, UNESCO y UKAT), y el código del concepto, en el 9,1% (AGROVOC).

Como se observa, en general los tesauros proporcionan bastante información de la dictada por la norma en los registros simples, por lo que es una visualización que han desarrollado bastante.

3.1.3 Visualización alfabética

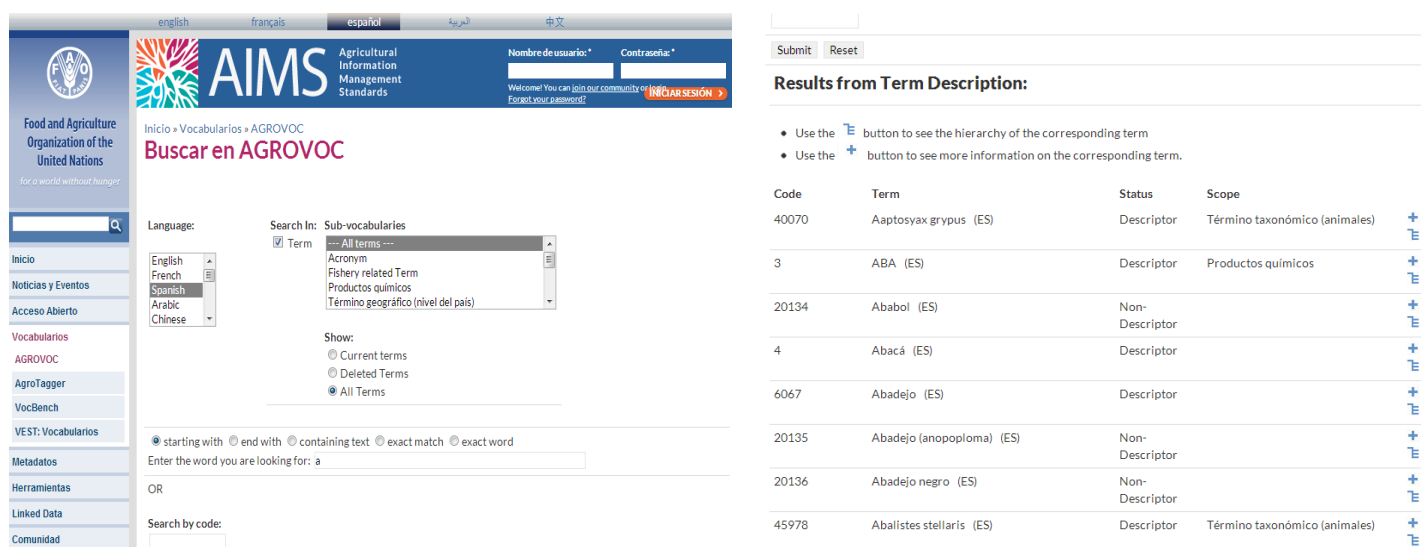
El caso de la visualización alfabética es especial y complejo de explicar, ya que los tesauros son muy heterogéneos entre ellos en cuanto a este tipo de presentación, por lo que es necesario comentar los acuerdos a los que hemos llegado.

Como se aprecia en la tabla 1, siete tesauros tienen una presentación alfabética. Pero, ¿qué hemos considerado visualización alfabética?: aquellos tesauros que tienen una opción en la pantalla principal que permite explorar una lista alfabética por letras, es decir, un listado completo de todos los términos que empiezan por una letra determinada.

The screenshot shows the website of the 'INSTITUTO DE ESTUDIOS DOCUMENTALES SOBRE CIENCIA Y TECNOLOGÍA (antes CINDOC) Consejo Superior de Investigaciones Científicas'. The main navigation bar includes 'Contacto', 'Tarifas', a search box with a 'Buscar' button, and a language selector for 'Iniciio' and 'English'. Below the navigation bar, there are tabs for 'Tesauros', 'Inicio', 'Jerárquico', 'Alfabético', and 'Búsquedas'. The 'Alfabético' tab is selected, displaying the 'Tesauro de Biblioteconomía y Documentación'. The page features a sidebar on the left with categories like 'EI IEDCYT', 'Investigación', 'Productos', and 'Acceso Gratuito'. The main content area shows an alphabetical list of terms, with the first letter 'A' highlighted. The list includes terms such as 'Accesibilidad', 'Accesibilidad de la información', 'Acceso a la información', 'Acceso al documento archivístico', 'Acceso al documento bibliotecario', 'Acceso al documento electrónico', 'Acceso gratuito', 'Actas', 'Actas de congresos', 'Actividad museística', 'Actividad profesional', 'Actividades culturales archivísticas', 'Actividades culturales bibliotecarias', 'Actualización profesional', 'Administración de museos', and 'Administración de personal'.

Imagen 10. Ejemplo del tesauro de Biblioteconomía del CINDOC con visualización alfabética

En los otros cuatro, es necesario que el usuario tome la iniciativa de iniciar una búsqueda en el “cajón de búsquedas” para poder visualizar alguna parte del listado alfabético. En tres tesauros (UNESCO, AGROVOC y Thesaurus for the Social Sciences), la opción por defecto permite que con teclear una letra sea suficiente para que se listen alfabéticamente todos los descriptores que empiezan por...

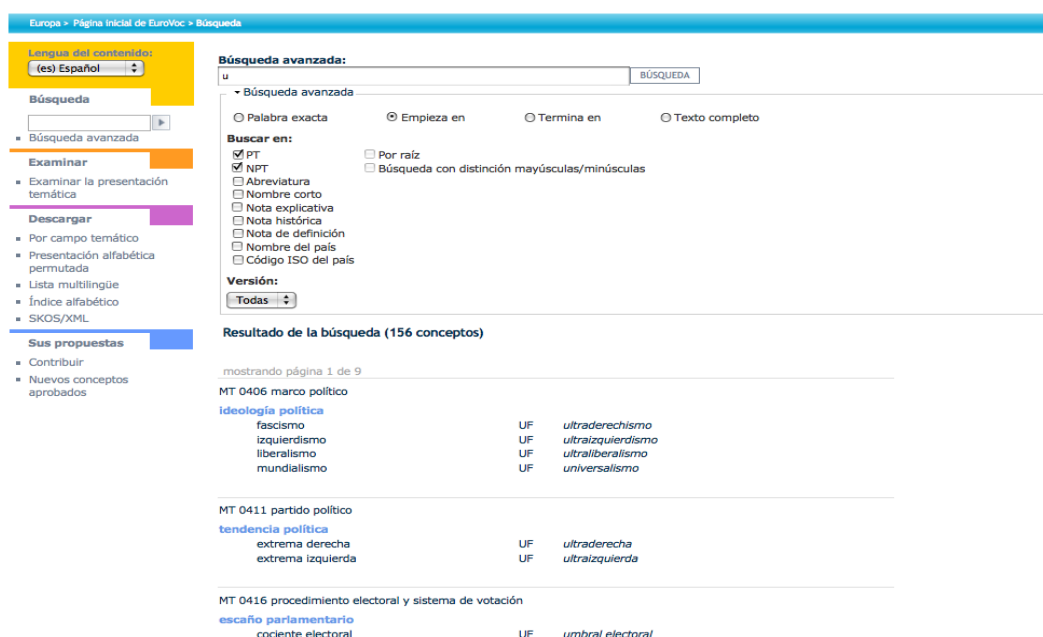


The screenshot shows the AGROVOC search interface. On the left, there's a sidebar with navigation links like 'Inicio', 'Noticias y Eventos', 'Acceso Abierto', 'Vocabularios', 'AGROVOC', 'AgroTagger', 'VocBench', 'VEST: Vocabularios', 'Metadatos', 'Herramientas', 'Linked Data', and 'Comunidad'. The main area has a search bar and a 'Buscar en AGROVOC' button. Below the search bar, there are options for 'Language' (English, French, Spanish, Arabic, Chinese) and 'Search In' (Term, Sub-vocabularies). The 'Show' options are 'Current terms', 'Deleted Terms', and 'All Terms'. The search criteria are 'starting with', 'end with', 'containing text', 'exact match', and 'exact word'. The search results are displayed in a table with columns: Code, Term, Status, and Scope.

Code	Term	Status	Scope
40070	Aptosyax grypus (ES)	Descriptor	Término taxonómico (animales)
3	ABA (ES)	Descriptor	Productos químicos
20134	Ababol (ES)	Non-Descriptor	
4	Abacá (ES)	Descriptor	
6067	Abadejo (ES)	Descriptor	
20135	Abadejo (anoploploma) (ES)	Non-Descriptor	
20136	Abadejo negro (ES)	Non-Descriptor	
45978	Abalíes stellaris (ES)	Descriptor	Término taxonómico (animales)

Imagen 11. Ejemplo del tesoro AGROVOC: búsqueda por “a”

En el caso de EUROVOC, la búsqueda por defecto obliga a que coincida exactamente el término tecleado con todo o parte del descriptor. Para encontrar una opción similar a la que hemos expuesto en el párrafo anterior, el usuario tiene que recurrir a la opción de búsqueda avanzada.



The screenshot shows the EUROVOC search interface. On the left, there's a sidebar with navigation links like 'Europa', 'Página inicial de EuroVoc', 'Búsqueda', 'Lengua del contenido', 'Búsqueda', 'Examinar', 'Descargar', 'Sus propuestas', and 'Contribuir'. The main area has a search bar and a 'Búsqueda avanzada' button. Below the search bar, there are options for 'Palabra exacta', 'Empezar en', 'Termina en', and 'Texto completo'. The 'Buscar en' options are 'PT', 'Por raíz', and 'Búsqueda con distinción mayúsculas/minúsculas'. The 'Versión' is 'Todas'. The search results are displayed in a table with columns: MT, Ideología política, and UF.

MT	Ideología política	UF
MT 0406 marco político	fascismo	UF ultraderechismo
	izquierdismo	UF ultraiquierdismo
	liberalismo	UF ultraliberalismo
	mundialismo	UF universalismo
MT 0411 partido político	extrema derecha	UF ultraderecha
	extrema izquierda	UF ultraiquierda
MT 0416 procedimiento electoral y sistema de votación	escanío parlamentario	UF umbral electoral

Imagen 12. Ejemplo de búsqueda avanzada en el tesoro EUROVOC: búsqueda por “u”

3.1.3.1 Características de la visualización alfabética

Una vez que hemos aclarado el concepto de visualización alfabética, pasamos a ver las características que deben tener los tesauros que cuenten con esta visualización. Según la norma ISO 25964-1 son:

- Los términos (preferentes y no preferentes) están ordenados en una secuencia alfabética única
- Los términos (preferentes y no preferentes) se distinguen tipográficamente
- Se incluye el registro completo de cada término como si fueran registros simples
- Se muestran los términos genéricos y específicos más inmediatos
- Se utiliza la sangría para los términos subordinados
- Los términos (preferentes y no preferentes) se presentan en un índice permutado

De los siete tesauros que tienen una opción de listado alfabético como tal, los tesauros del proyecto UKAT, GCC y CINDOC tienen los términos preferentes y no preferentes ordenados en una secuencia alfabética única y en los dos últimos, los términos se distinguen tipográficamente. El resto de características no las cumple ningún tesoro.

Como se observa, en general esta presentación está muy presente en los tesauros, lo que facilita a los usuarios la tarea de navegación y recuperación de información.

3.1.4 Visualización jerárquica

Encontramos una visualización jerárquica en 10 tesauros de los 11 que hemos analizado (90,9%). El único que no cuenta con ella es Reegle que, como ya hemos indicado, a pesar de que en su nombre oficial contiene la palabra tesoro (y por eso fue elegido), en realidad es un glosario que cuenta con una presentación alfabética pero con ningún tipo de presentación en forma de jerarquía.

Este tipo de visualización, en su primer nivel, pone de manifiesto la macroestructura del tesoro mediante una organización de los conceptos en temas o facetas. La norma le da mucha importancia a la organización de los conceptos de manera facetada. Sin embargo, de los 10 tesauros que tienen visualización jerárquica, y a pesar de que todos están preparados para funcionar en el entorno de la Web Semántica, el 90% organiza los conceptos por temas/áreas/temáticas/disciplinas, y solo uno, Social Semantic Web, por facetas. Una posible causa es que, tradicionalmente en el ámbito de las herramientas de organización del conocimiento (véase taxonomías, clasificaciones, organización de clasificaciones bibliográficas), las facetas nunca han sido muy utilizadas. Además, no es un método muy familiar para los usuarios.

3.1.4.1 Características de la visualización jerárquica

Las características que deben tener los tesauros que cuenten con esta visualización, según la norma ISO 25964-1 son:

- Existe una lista de términos generales que no se asocia con ninguna categoría sino que son válidos para todas
- Aparece el término general seguido del resto de términos con distintos niveles de sangría
- Los términos están organizados en facetas dentro de los campos temáticos

La existencia de una lista de términos generales en un tesoro está muy relacionada con la organización de los términos en facetas. Solo un tesoro, STW Economics la tiene, aunque este hecho puede resultar extraño ya que este tesoro no organiza sus conceptos por facetas, sino por temas. Ya hemos comentado que solo Social Semantic Web los organiza mediante una clasificación facetada.

Los distintos niveles de sangría aparecen en seis de los 10 tesauros (60%); no aparecen en el Social Semantic Web, el GCC, el STW Economics y en GEMET. Consideramos que el sangrado es básico para diferenciar los distintos niveles que existen dentro de los elementos, y, de esta forma, la jerarquía que hay entre los términos. En algunos casos podría contribuir a diferenciar términos generales de términos relacionados. En los otros casos, los conceptos se listan todos unos debajo de otros.

Por último, la organización de facetas dentro de los campos temáticos no se cumple en ninguno de los 10 tesauros (0%). Ya hemos explicado que no es un tipo de organización muy tradicional dentro del ámbito de las herramientas de organización del conocimiento, a pesar de que se intente fomentar su uso.

3.1.4.2 Información adicional en la jerarquía

Uno de las características de esta visualización es la información adicional que se muestra de cada término. Siete de los diez tesauros que tienen visualización alfabética no incluyen información adicional en la jerarquía de la que indica la norma, (puede incluir otro tipo de información). Esto indica que solo el 30% de las herramientas se ajustan a la norma respecto a este tema. Pero, ¿qué información adicional dice la norma que se debería incluir en los registros?

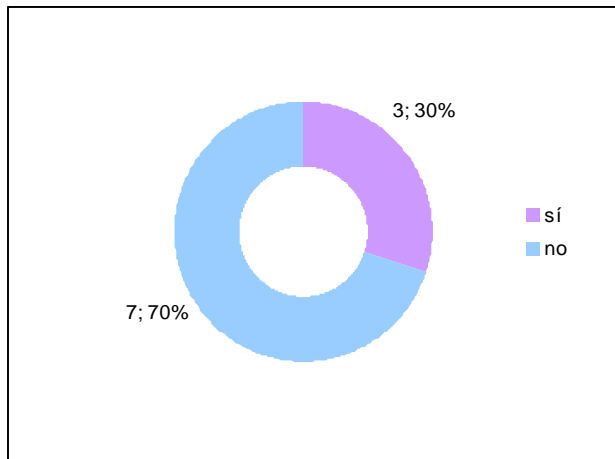


Gráfico 2. ¿Cuántos gráficos tienen información adicional?

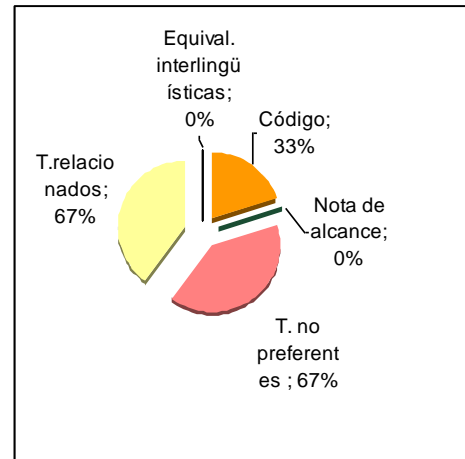


Gráfico 3. ¿Qué información adicional contienen?

De los tres tesauros que incluyen alguna información adicional, dos incluyen los términos no preferentes (67%), que son AGROVOC y UNESCO. También en un 67% (EUROVOC y AGROVOC) se ven las relaciones asociativas. Además, sólo AGROVOC incluye también el código del concepto y ninguno muestra las equivalencias interlingüísticas ni las notas de alcance.

Como se observa, AGROVOC es el tesauro que mejor se ha adaptado a la normativa en este apartado, incluyendo tres de las cinco características establecidas por la norma.

En general, para estos resultados hay que tener en cuenta que algunos tesauros pueden incluir información que no está especificada en la norma. Es lógico que aparezcan los términos no preferentes y los relacionados por ser considerados elementos básicos y que no tengan cabida las equivalencias interlingüísticas ni las notas de alcance, ya que si no parecería que una visualización de registro simple mas que una visualización jerárquica.

El balance general de esta presentación es positivo aunque es necesario que los tesauros mejoren en cuanto a la inclusión de algún término más en la jerarquía y la mejora de algunas características, como el sangrado.

3.1.5 Visualización gráfica

La norma ISO 25964-1 solo menciona esta visualización en un pequeño apartado diciendo que es compleja de crear y de actualizar y que se estropeará si incluye demasiada información adicional. Sin embargo, consideramos que es una presentación que complementa muy bien al resto de visualizaciones, facilitando al usuario la comprensión de las relaciones establecidas entre los conceptos.

Los resultados de este tipo de visualización son francamente negativos. De los 11 tesauros analizados, solo tres (27,3%), EUROVOC, Social Semantic Web y Reegle, cuentan con esta visualización. Los tres pivotan sobre un concepto como elemento principal, mostrando sus

relaciones con otros conceptos. Se puede navegar y entre los términos del tesauro; si pinchas en una visualización cambia y pivota sobre ese nuevo término.

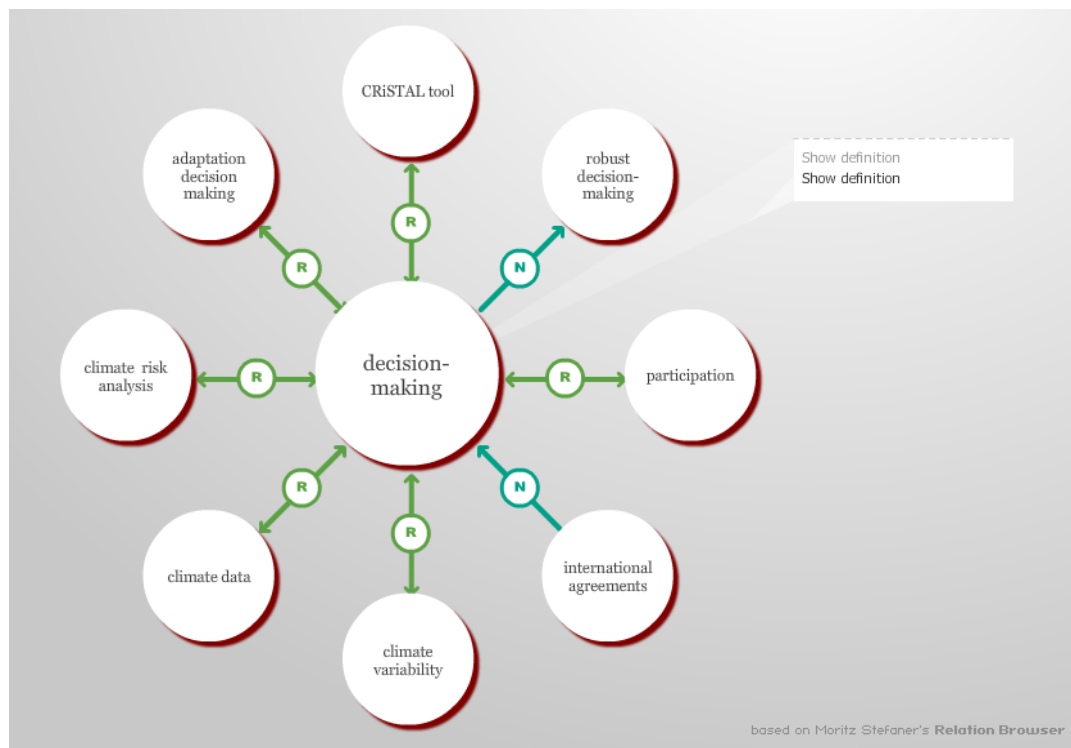


Imagen 13. Visualización gráfica del concepto “decision-making” en el tesauro Reegle

3.1.6 Presentaciones multilingües

Ésta es la última visualización que indica la norma que deben tener los tesauros electrónicos. Solo aparece en siete de los once tesauros (63,6%).

Ya se comentó en el apartado “3.1 Tipos de visualizaciones”, que los datos obtenidos con respecto a esta presentación son negativos porque todos los tesauros debería ofrecer esta presentación, junto con el registro simple, ya que es básico que hoy en día los tesauros faciliten su consulta a los usuarios proporcionándoles la posibilidad de consultarlos en varios idiomas.

3.1.6.1 Características de las visualizaciones multilingües:

Según la ISO 25964-1, las características que debe tener una visualización multilingüe son:

- Se permite la visualización simultánea en dos idiomas
- Se permite cambiar el idioma de la pantalla
- Cada término preferente tiene su equivalencia interlingüística
- Cada idioma se identifica con un código siguiendo la norma ISO 639-1
- La secuencia alfabética es distinta para cada idioma

- Las equivalencias interlingüísticas se muestran en una tabla de correspondencia independiente

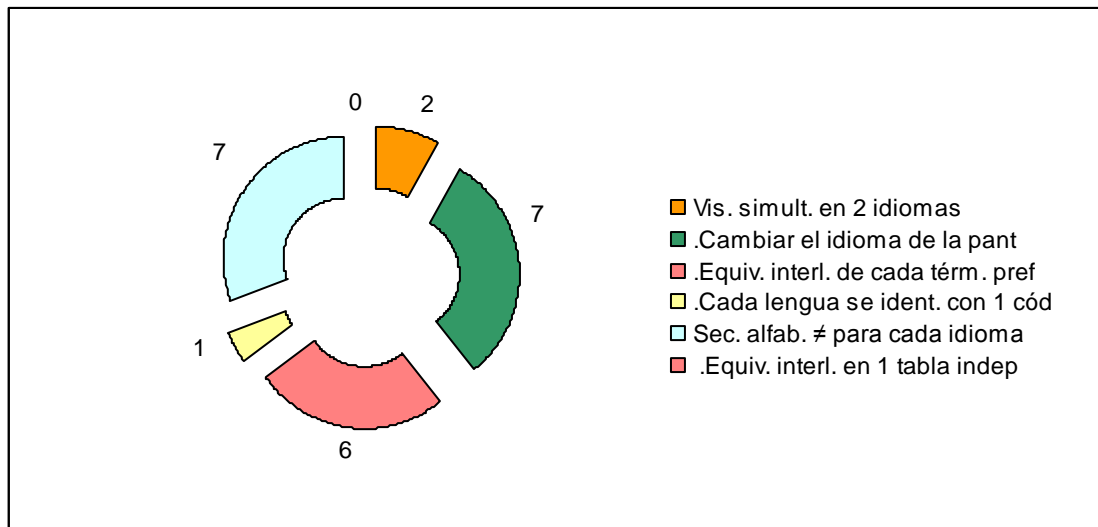


Gráfico 4. Características de la visualización multilingüe

En todos los tesauros que tienen este tipo de visualización (7), se puede cambiar el idioma de la pantalla y la secuencia alfabética es distinta para cada idioma. En todos menos en uno (85,7%), que es AGROVOC, cada término preferente tiene su equivalencia interlingüística en los idiomas en los que se puede consultar el registro simple. Esta característica facilita cambiar el idioma de consulta de la herramienta de manera fácil y rápida, únicamente pinchando en el término extranjero deseado (que se encontrará debajo del término preferente o al lado). En el caso de AGROVOC, es necesario pinchar en el idioma en el que se desea visualizar el concepto para cambiar toda la visualización. Las opciones de idiomas suelen situarse en la parte superior del tesoro.

La visualización simultánea en dos idiomas sólo está disponible en el tesoro de Ciencias Sociales y en el de GEMET (28,6%). Consideramos que esta opción es muy interesante aunque somos conscientes de que requiere un esfuerzo mayor por parte de las instituciones que gestionan los tesauros. Con la visualización simultánea en dos idiomas el usuario podrá comparar más fácilmente las relaciones de un tesoro en inglés y en su lengua materna, por ejemplo. Quizá de este dato podemos extraer la conclusión positiva de que, aunque solo dos tesauros dan esta opción, en los siete que cuentan con esta visualización se puede cambiar el idioma de la pantalla de una manera o de otra, por lo que es posible visualizar el registro en otro idioma de manera fácil.

Por otro lado, solo EUROVOC (14,3%) identifica las lenguas con un código de dos letras siguiendo la norma ISO 639-1. La explicación que podemos ofrecer es que este tesoro es el segundo que más idiomas ofrece a los usuarios (por detrás de GEMET, que tiene 33), para poder

consultarlo, hasta un total de 24, por lo que tiene que identificar muy bien cada uno de ellos para facilitar la navegación a los usuarios.

Por último, ningún tesoro tiene sus equivalencias interlingüísticas en una tabla de correspondencia independiente; en general, los tesauros poseen la equivalencia de los términos preferentes debajo del término o al lado, pero en ningún caso en una zona independiente, siempre se encuentran dentro de la misma página.

3.1.7 Gráfico general del apartado de visualizaciones por tesauros

A continuación mostraremos un gráfico en el que se puede comparar la valoración total de los tesauros en el apartado de visualización. En este apartado se valoraban 35 características.

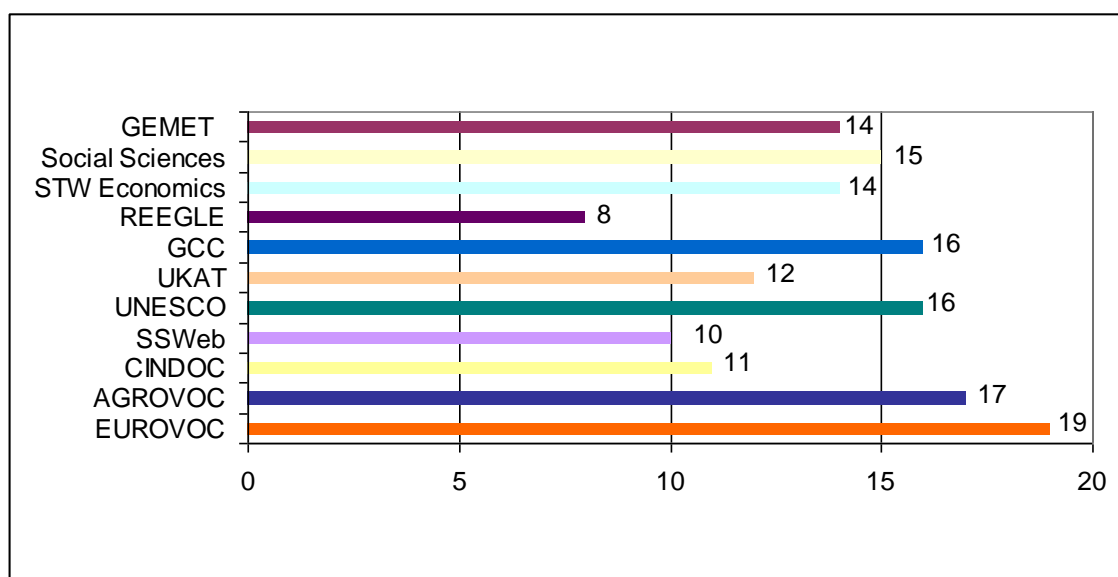


Gráfico. 5. Características de la visualización por tesauros

El mejor tesoro en cuanto a visualización es EUROVOC, que incluye el 54,3% de las características analizadas. Este tesoro destaca principalmente por la cantidad de elementos que contienen el registro simple y por las características que presenta en las visualizaciones multilingües. No tiene la opción de visualización alfabética pero sí gráfica.

Le sigue de cerca con 17 características, AGROVOC, obteniendo un resultado positivo en el 48,6% de las variables, algo menos de la mitad. Destaca también por la cantidad de elementos que contienen los registros simples y por la información adicional que se presenta en la visualización jerárquica. No tiene presentación gráfica, las presentaciones multilingües son un poco pobres y para acceder a la presentación alfabética es necesario que el usuario tome la iniciativa escribiendo una letra.

A continuación están el tesoro de la UNESCO y del Government of Canada Core Subject, con 16 características conseguidas (45,7%), Ciencias Sociales con 15 (42,9%) y STW Economics y GEMET con 14 variables (40%). Esta situación se justifica por los buenos resultados que proporciona la gran cantidad de elementos que tienen los registros simples, aunque cada uno falla en otras visualizaciones o características.

Con 12 variables está el tesoro del UKAT (34,3%), con 11 el CINDOC (31,4%) y con 10 el Social Semantic Web. La inexistencia de presentaciones multilingües o de información adicional en la visualización jerárquica son las principales carencias de estos tesauros. El último lugar en cuanto a visualización lo ocupa Reegle, que al ser un glosario no sigue las pautas de la ISO 25964-1, no teniendo visualización jerárquica ni multilingüe.

Para finalizar el tema de las visualizaciones, consideramos que los resultados son relativamente positivos. Sin embargo, es necesario que estas herramientas se esfuercen por cumplir alguna pauta más de la norma. Esto ayudaría a mejorar un aspecto tan básico para los usuarios como son las distintas maneras en las que pueden consultar los registros que contiene un tesoro.

3.2 Búsqueda

Las opciones de búsqueda que proporciona el tesoro son un aspecto fundamental para facilitar el acceso de los usuarios a la información. Según la ISO 25964-1 deberían ser:

- Posibilidad de buscar por términos que contienen cualquier palabra
- Posibilidad de buscar por términos que contienen una combinación de palabras
- Posibilidad de buscar por términos que contienen una combinación de caracteres
- Posibilidad de buscar por términos y conceptos utilizando el código que los identifica
- Posibilidad de buscar por términos equivalentes en el mismo idioma
- Posibilidad de buscar por términos equivalentes en distinto idioma

Los 11 tesauros que hemos analizado tienen cajón de búsqueda pero en el tesoro de Social Semantic Web no funciona porque no busca por ninguna palabra.

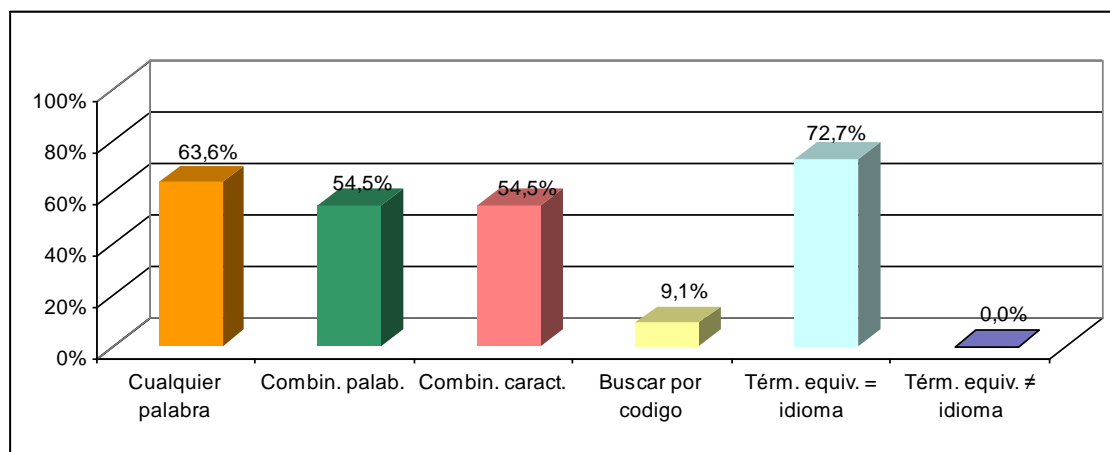


Gráfico 6. Opciones de búsqueda

Lo primero que observamos con estos resultados es que el simple hecho de tener una opción de búsqueda no garantiza buscar por un índice permutado. La opción que más cumplen los tesauros, hasta un total de 8 (72,7%), es la posibilidad de buscar términos equivalentes en la misma lengua. Este hecho es muy positivo, pues significa que los tesauros tienen un buen sistema interior de reenvíos entre términos y conceptos. Sin embargo, y a pesar de que, como ya hemos comentado, el 63% de los tesauros (7 de los 11) ofrece la posibilidad de visualizaciones multilingües, la búsqueda entre términos equivalentes en diferentes lenguas no la ofrece ninguna herramienta. Consideramos que si se da la posibilidad de consultar el tesoro en otros idiomas sería muy recomendable habilitar también esta opción.

Como ya hemos comentado también, AGROVOC es el único tesoro que asigna un código a cada uno de los conceptos. Aplicando esto a las opciones de búsqueda que estamos analizando ahora, este instrumento es, lógicamente, el único que permite buscar por códigos, algo que personalmente a mí me ha facilitado mucho la consulta del tesoro cuando estaba buscando ejemplos para analizar sus características. En este caso, consideramos que no sería tan complejo asignar un código único y personal a cada concepto para que se identificaran inequívocamente (además de la URI que tienen asignada al estar etiquetados con SKOS), y sería un hecho muy agradecido por parte de los usuarios ya que les facilitaría el browsing por el tesoro. Además, los códigos representan la estructura sistemática de la herramienta y contextualizan al descriptor en un área temática.

Por otro lado, y respecto a las tres primeras opciones de búsqueda relacionadas con el tipo de términos que los tesauros permiten buscar, siete tesauros (63,6%) permiten buscar por palabras clave dentro del descriptor, seis (54,5%) permiten combinar palabras en la búsqueda, y seis también permiten buscar a partir de una combinación de caracteres. De los dos últimos datos, cinco tesauros coinciden en permiten buscar por una combinación de palabras y a partir de caracteres, que son EUROVOC, AGROVOC, CINDOC, UNESCO y UKAT.

Por último, y relacionando estos resultados con los obtenidos en el apartado de visualización, de los cuatro tesauros en los que el usuario tenía que tomar la iniciativa para poder visualizar un orden alfabético de los términos, uno, el de Ciencias Sociales, tampoco permite buscar por palabras clave dentro del descriptor.

A nivel general, los tesauros cumplen a medias las características marcadas por la norma ISO 25964-1. La parte de la búsqueda es fundamental en un tesoro, si es complicada, los usuarios abandonarán el tesoro porque les resultará complicado obtener la información que desean. Las instituciones tienen que poner más empeño en mejorarla.

A continuación mostraremos un gráfico que compara la valoración de los tesauros en el apartado de búsqueda. Las características que se analizaban en este apartado son 6.

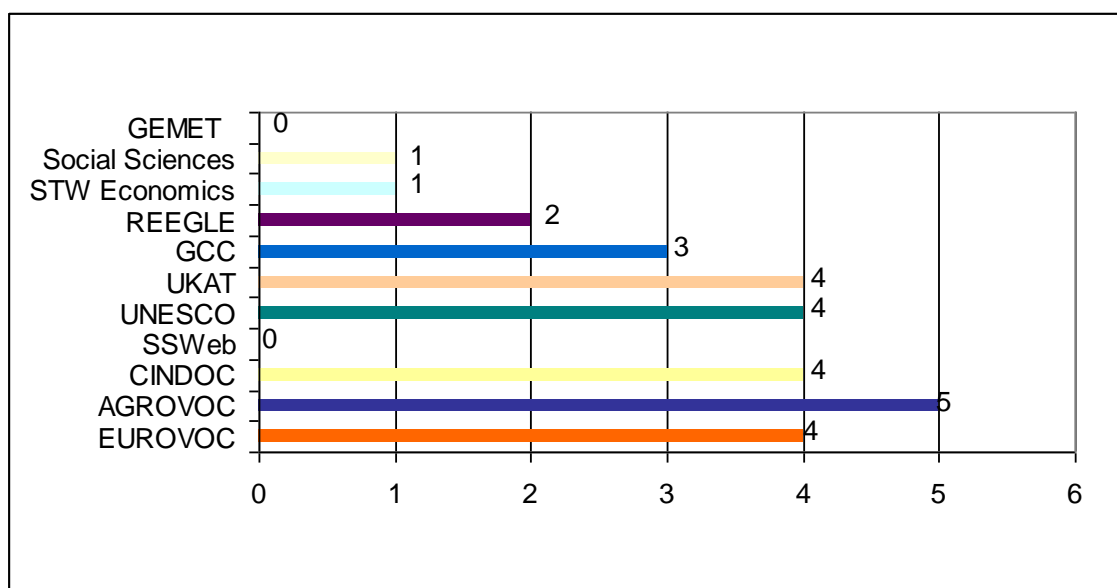


Gráfico 7. Características de la búsqueda por tesauros

El mejor tesoro en cuanto a búsqueda es AGROVOC, que incluye el 83,3% de las características analizadas. Destaca principalmente por la posibilidad de buscar por código y por cumplir todos los requisitos de búsqueda posibles (cualquier palabra, combinación de palabras y combinación de caracteres). Le siguen de cerca EUROVOC, CINDOC, UNESCO y UKAT con cuatro variables conseguidas (66,7%). Estos tesauros cumplen exactamente los mismos requisitos, no contando únicamente con la posibilidad de buscar términos equivalentes en otro idioma.

Con tres características (50%) encontramos el tesoro del GCC y con dos (33,3%), el Reegle. Respecto al primero, carece de la posibilidad de buscar por términos que contienen una combinación de caracteres y de buscar por términos interlingüísticos; Reegle, en cambio, buscar por términos que contienen cualquier palabra o una combinación de caracteres pero no incluye las demás opciones. Es necesario que ambos tesauros mejoren sus algunas de estas opciones de búsqueda para poder conseguir más usuarios que consulten su tesoro.

Los últimos puestos los ocupan STW Economics y Ciencias Sociales (con una sola característica conseguida (16,7%), permitiendo ambos solo buscar términos equivalentes en el mismo idioma; y Social Semantic Web y GEMET que no cumplen ninguna característica de la norma en cuanto a la búsqueda. Estos cuatro tesauros tienen que replantearse completamente la reconstrucción de las opciones de búsqueda para su tesoro porque es una opción básica para su funcionamiento y consulta.

Para finalizar el tema de la búsqueda es necesario decir que 5 tesauros (45,4%) no cumplen con la mitad de las propuestas de la normativa y que es necesario que la reestructuren para facilitar a los usuarios esta tarea y para ofrecer un buen servicio de consulta.

3.3 Navegación

Las opciones de navegación que la norma ISO 25964-1 indica que se deben cumplir son:

- Posibilidad de moverse desde el registro de un concepto hacia sus términos más genéricos
- Posibilidad de moverse desde el registro de un concepto hacia sus términos más específicos
- Posibilidad de contraer y expandir los niveles de las visualizaciones jerárquicas/sistemáticas
- Posibilidad de cambiar de un tipo de visualización a otra
- Posibilidad de ver dos visualizaciones de forma simultánea

Los resultados de este apartado en solitario son francamente los mejores de todas las variables que hemos analizado. De los 11 tesauros a los que hemos pasado el cuestionario, todos ellos tienen la posibilidad de navegar hacia los términos más genéricos y específicos del concepto seleccionado y cambiar de un tipo de visualización a otra.

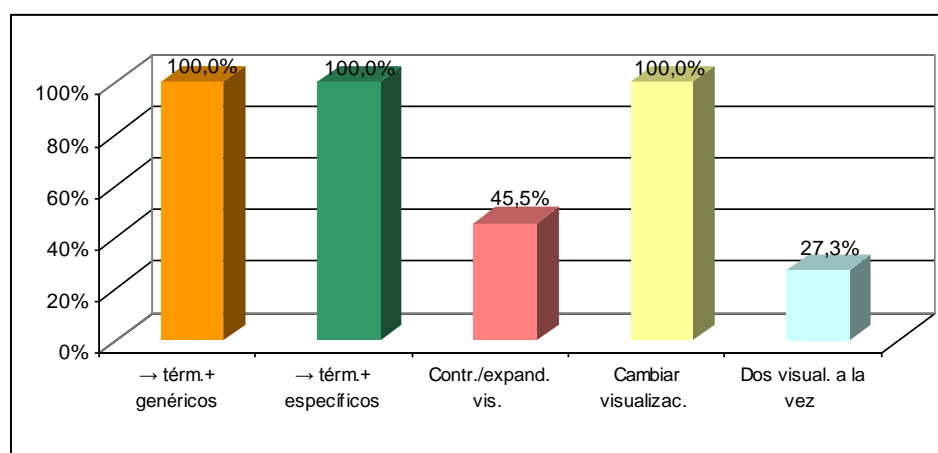


Gráfico 8. Opciones de navegación

La contracción y expansión de los niveles en las visualizaciones jerárquicas es una opción presente en cinco de los once tesauros (45,5%). Quizá no es un dato muy importante, pues lo verdaderamente relevantes es la posibilidad de realizar las acciones anteriores para que el usuario no tenga que volver a empezar una búsqueda si quiere consultar conceptos cercanos.

Por último, somos conscientes de que la visualización simultánea de dos presentaciones es compleja, pero no estamos especialmente de acuerdo con esta recomendación de la norma ya que somos más partidarios de que las visualizaciones se complementen. A la vista de los resultados, los tesauros tampoco están muy de acuerdo con ella, ya que solo se cumple en el 27,3%.

A nivel de opciones de navegación los resultados han mejorado considerablemente con respecto a las opciones de búsqueda. A continuación mostraremos un gráfico que compara las variables que han conseguido los tesauros en el apartado de navegación. Las características que se analizaban en este apartado son 6.

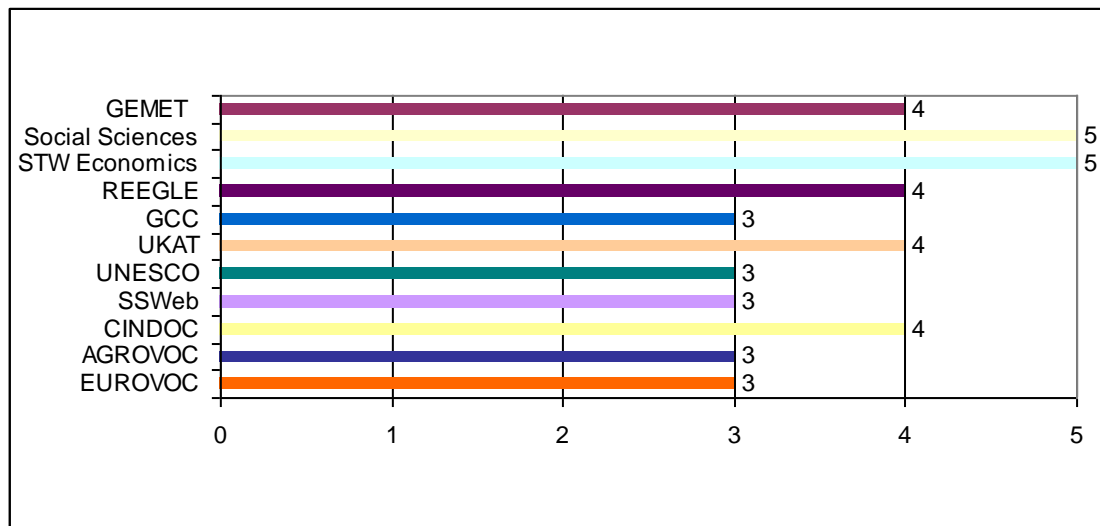


Gráfico 9. Características de la navegación por tesauros

Dando la vuelta a los resultados obtenidos en la búsqueda, STW Economics y Ciencias Sociales ocupan los primeros puestos en cuanto a navegación, consiguiendo el total de las características (100%). Les siguen de cerca CINDOC, UKAT, GEMET y Reegle, con cuatro características (80%). A los tres primeros solo les faltaría cumplir la opción de visualizar simultáneamente dos presentaciones y a Reegle, que sí cumple esa recomendación, la de expandir y contraer los niveles en las visualizaciones jerárquicas. Con tres características (60%) se quedan el resto de tesauros, coincidiendo principalmente en que no cuentan con las visualizaciones simultáneas.

La verdad es que analizando uno por uno los resultados de los tesauros de forma individual no parecía que cumplieran muchas de las características pero viéndolo en conjunto es justo decir que

los tesauros han desarrollado de forma notable este apartado y han conseguido tener un sistema de navegación fácil para los usuarios.

3.4 Búsqueda y navegación

A continuación mostraremos en un gráfico los resultados acumulativos en cuanto a búsqueda y navegación que han obtenido los tesauros. Recordemos que en este apartado se analizaban 11 características: 6 en la búsqueda y 5 en la navegación.

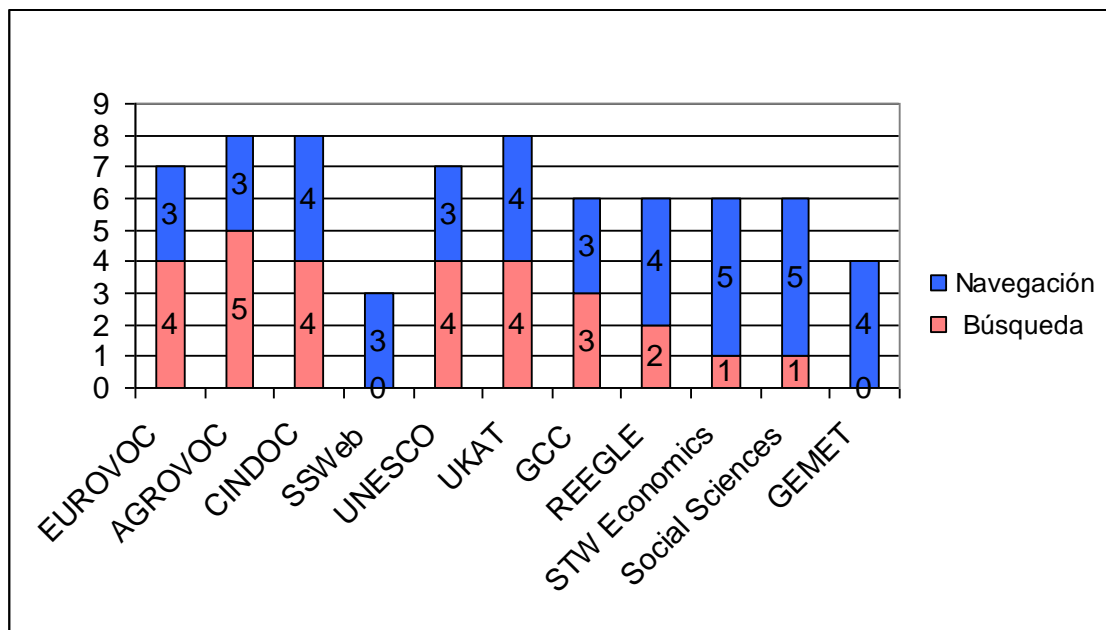


Gráfico 10. Características de la búsqueda y navegación por tesauros.

Como ya hemos comentado, en general los tesauros han conseguido más características gracias a los buenos resultados obtenidos en el apartado de navegación.

Tres tesauros (AGROVOC, CINDOC y UKAT) cumplen 8 características (72,7%) ajustándose, por lo tanto, bastante bien las recomendaciones dictadas por la norma. Además, prácticamente cumplen el mismo número de características en los dos apartados.

Los siguen de cerca EUROVOC y UNESCO con 7 puntos (63,6%), que casualmente cumplen más objetivos de la parte de búsqueda que de navegación, y Government of Canada Core Subject, Reegle, STW Economics y Ciencias Sociales con 6 (54,5%), mayoritariamente del apartado de navegación. Todos ellos han cumplido la mitad de lo expuesto por la normativa por lo que se pueden considerar resultados satisfactorios.

En último lugar se encuentran Social Semantic Web y GEMET; en ambos tesauros las características que han obtenido, que son 3 (27,3%) y 4 (36,4%) respectivamente han sido gracias

a los resultados obtenidos en el apartado de navegación. En estos dos tesauros la norma no se cumple de manera notable en cuanto a la búsqueda, por lo que es necesario que la revisen e intenten aplicarla con más profundidad.

Para finalizar el apartado de la navegación y la búsqueda en el tesoro, volvemos a remarcar los buenos resultados en cuanto a la primera opción, y los bastante negativos en la segunda. Algunas herramientas necesitan una revisión en profundidad de la normativa para poder satisfacer mejor las necesidades de los usuarios.

3.5 Integración con otras aplicaciones

La última parte de nuestra investigación gira entorno a la posibilidad de que los tesauros se integren en otras aplicaciones y se les de un sentido más práctico de búsqueda de información, no solo de normalización y de consulta.

Lo primero que cabe destacar es que solo tres de los once tesauros (27%), UNESCO, Reegle y STW Economics, están integrados en algún tipo de aplicación desde su página principal.

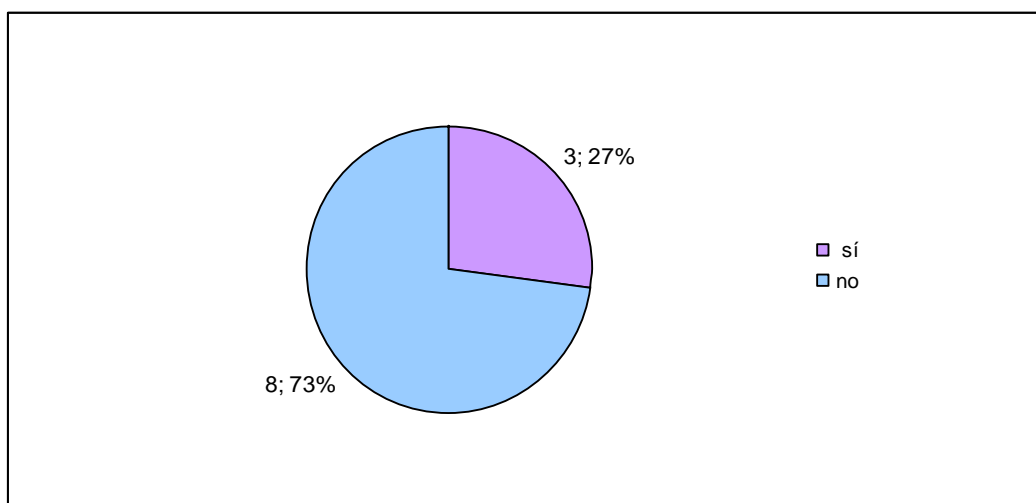


Gráfico 11. Integración

De las características que indica la norma ISO 25964-1, solo se cumplen cuatro (36,4%) por parte de algún tesoro.

- El tesoro da la opción de seleccionar un término que se utiliza en una búsqueda
- El tesoro da la opción de seleccionar más de un término que se utiliza en una búsqueda
- El tesoro tiene la opción de construir una estrategia de búsqueda copiando términos desde el tesoro
- El tesoro tiene la opción de “explode”

- En la búsqueda, ¿se sustituye automáticamente un término no preferente por uno preferente?
- En el proceso de especificación de la búsqueda, ¿el tesauro proporciona el número de publicaciones que tiene cada subconjunto de la colección?
- Cuando un usuario introduce un término similar a uno o más términos del tesauro, ¿se responde con “¿Quiso decir...?”, de manera que se ofrece al usuario coincidencias cercanas de términos que se encuentran en el tesauro?
- Cuando un usuario introduce un término similar a uno o más términos del tesauro, ¿éste responde recuperando los resultados de todos los términos y agrupándolos?
- ¿El usuario puede delimitar su búsqueda?
- Si la búsqueda da pocos resultados, ¿se puede expandir automáticamente a otros términos? (generales, específicos, relacionados...)
- Cuando un término de búsqueda se transforma, sustituye, traduce o reemplaza por otro, ¿se informa al usuario de este hecho?

Los tres tesauros arriba indicados (100%) tienen la opción de seleccionar un término que se utiliza en una búsqueda, dos (66,7%) pueden delimitar su búsqueda, y dos también puede expandirla a otros términos. El trío de tesauros permite acceder desde el descriptor a la base de datos pero una vez dentro, no permite crear ninguna estrategia de búsqueda; son sistemas estáticos. El resto de características no las cumple ningún tesauro.

A nivel de integración, a los tesauros les queda mucho camino por recorrer: primero, consiguiendo que se integren en otras aplicaciones, y segundo, cumpliendo lo establecido por la norma en cuanto a las características.

A continuación mostraremos un gráfico que compara el número de características que han obtenido los tesauros en el apartado de integración. Las características totales que podían conseguir en este apartado son 11.

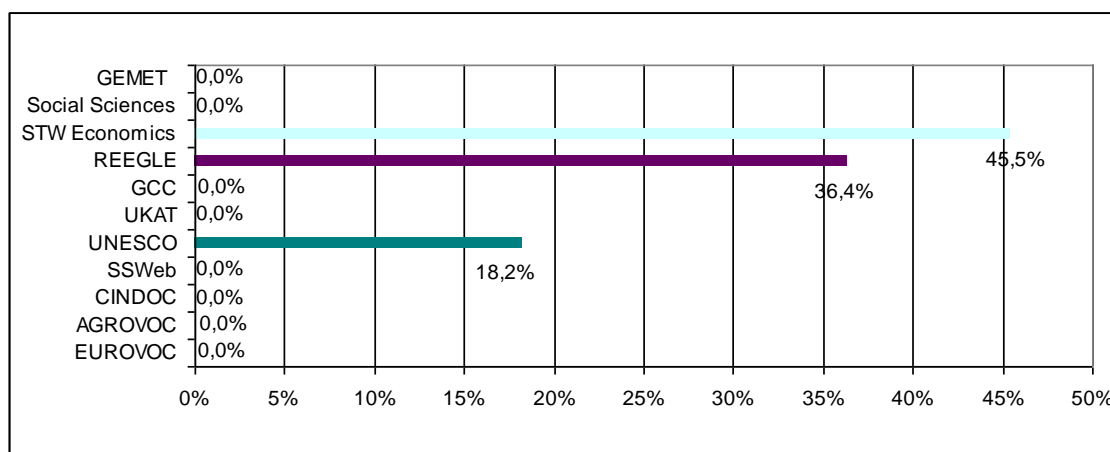


Gráfico 12. Características de la integración por tesauros

STW Economics consigue la primera posición con 5 características (45,5%): da la opción de seleccionar un término que se utiliza en una búsqueda, en el proceso de refinamiento proporciona el número de publicaciones y el usuario puede delimitar su búsqueda y expandirla automáticamente a otros resultados. Aunque no cumple demasiadas características, es el mejor tesoro en este apartado.

Reegle también está integrado con una aplicación, dando la opción de seleccionar un término que se utiliza en una búsqueda, permitiendo la expansión automática a otros términos, etc. El caso de la UNESCO es más comprensible, pues está integrado con otras aplicaciones de la misma institución: biblioteca, archivo, etc. Es una pena que en tesauros tan importantes como EUROVOC o AGROVOC no se pueda acceder a otras aplicaciones desde su página principal. El reto más importante de estos tesauros será alcanzar esa integración directa.

3.6 Valoración general

Por último vamos a mostrar un gráfico que aglutina los resultados de cada tesoro en cada uno de los apartados que hemos analizado.

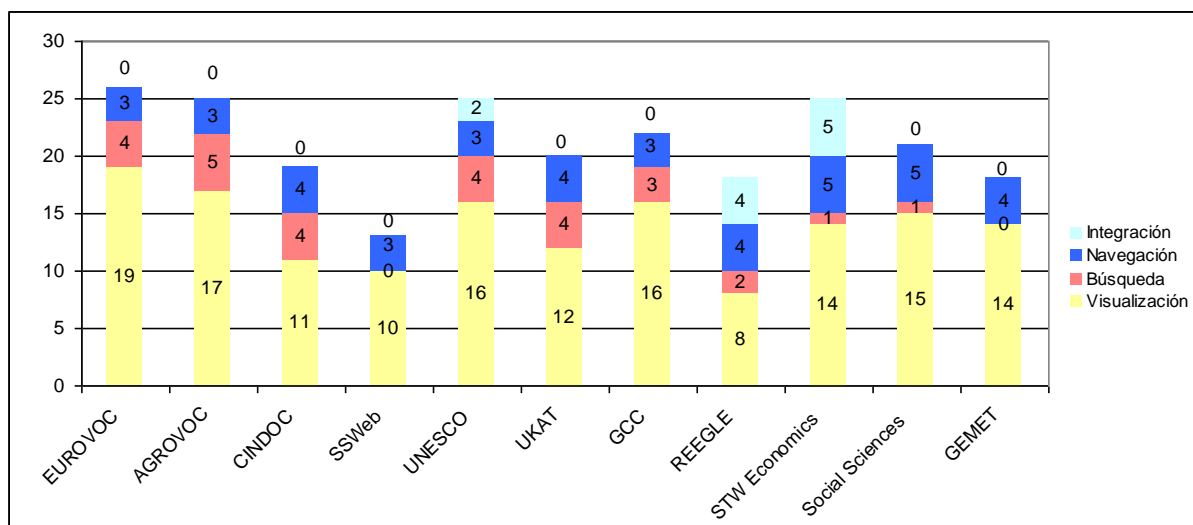


Gráfico 13. Valoración general

Para entender mejor los resultados añadiremos una tabla-resumen

	Visualización	Búsqueda	Navegación	Integración	Total	%
EUROVOC	19	4	3	0	26	45,6%
AGROVOC	17	5	3	0	25	43,9%
CINDOC	11	4	4	0	19	33,3%
Social Semantic Web	10	0	3	0	13	22,8%
UNESCO	16	4	3	2	25	43,9%
UKAT	12	4	4	0	20	35,1%
GCC	16	3	3	0	22	38,6%
REEGLE	8	2	4	4	18	31,6%
STW Economics	14	1	5	5	25	43,9%
Social Sciences	15	1	5	0	21	36,8%
GEMET	14	0	4	0	18	31,6%
<i>Total</i>	35 (61,4%)	6 (10,5%)	5 (8,8%)	11 (19,3%)	57	

Tabla 3. Valoración general

En verde se observa el tesoro que más características cumple de cada una de los apartados y en azul, las variables analizadas en cada apartado. A pesar de que STW Economics cumple más características en la navegación y en la integración, EUROVOC se postula como mejor tesoro, cumpliendo 26 características (45,6%), obtenidas gran parte de ellas en el apartado de visualizaciones. Le siguen de cerca AGROVOC, UNESCO y STW Economics, tesauros pertenecientes a grandes instituciones, con 25 características cada uno (43,9%).

La tabla muestra que aproximadamente un 61% de las características cumplidas por cada tesoro pertenecen al apartado de visualizaciones, que es del que más variables hemos analizado, hasta un total de 35. Le sigue la integración con 11 características, abarcando el 19% del total, la búsqueda con 6 características (10,5%), y la navegación con 5 (casi un 9%).

IV. CONCLUSIONES

En primer lugar es necesario destacar que no hay apartado de discusión por la imposibilidad de comparar este estudio con otros de características similares, ya que no hemos encontrado ningún trabajo afín a este para poder contrastar los resultados o, simplemente, para tener una orientación sobre qué criterio utilizar a la hora de seleccionar los tesauros que conformarían la muestra de nuestro trabajo. Este hecho también ha provocado que los resultados tengan tantos comentarios.

Evaluar los tesauros con la norma ISO 25964-1 ha sido muy complicado debido a que no es una herramienta muy clara. A nivel de apartados, el más difícil de valorar ha sido el de las visualizaciones (y dentro de ellas, la alfabética) donde la norma es muy ambigua y muestra una gran disparidad de ellas, sin establecer una diferenciación clara entre las imprescindibles y las que pueden ser complementarias. Esto nos ha dificultado mucho el trabajo, haciéndonos dudar muchas veces sobre la respuesta que teníamos que elegir en función de las variables que teníamos. En el aspecto de la búsqueda y navegación, en cambio, sí que da nociones más concisas sobre las características que tienen que cumplir los tesauros para facilitar estas operaciones a los usuarios. Por último, el apartado de la integración es el menos desarrollado, dando solo unas pautas generales (recordemos que la segunda parte de la norma se dedica en exclusiva a los formatos necesarios que tienen que tener estas herramientas para garantizar la interoperabilidad entre ellas y la propia integración), pero también más sencillas de evaluar y de aplicar que en el primer caso.

Como ya se ha explicado en la metodología, el trabajo ha comprendido varias fases, siendo la más compleja de todas, la valoración de las características en algunos tesauros. El problema ha surgido en el momento en el que una misma visualización es tratada de diferente manera en dos tesauros, como ocurre en el caso de la alfabética. Las variables del cuestionario, que, como ya hemos dicho, están extraídas de la norma ISO 25964-1, son muy claras y no dan opción a elegir un término medio, por lo que en ocasiones ha sido complicado elegir la respuesta adecuada para cada pregunta en cada tesoro.

Como se observa, no ha sido nada fácil extraer los datos de los diferentes tesauros. Cada tesoro “es un mundo”, y, aunque todos están creados a partir de unos mismos objetivos y con unos elementos básicos, cada uno hace después su propia interpretación de la norma y la adapta a sus necesidades. La norma ya está redactada; ahora es el turno de las instituciones de adaptarla a sus herramientas.

La realización de este trabajo ha supuesto para mí una mejora en el uso de algunas herramientas como son el análisis de una norma internacional, el manejo de Excel, la tabulación de resultados o el análisis de datos.

En el apartado de resultados, las conclusiones que hemos sacado de nuestro trabajo son:

1. En cuanto a la identificación de los tipos de visualizaciones, hemos comprobado que los tesauros, según la norma, deben contar con: un registro simple de cada concepto que contenga distintos tipos de términos y donde queden reflejadas las distintas relaciones que se pueden establecer entre ellos (el 100% de los tesauros la tienen; respecto a los elementos, todos los tesauros contienen la mayoría de los dictados por la norma, excepto el código del concepto, que solo lo incluye AGROVOC), una presentación alfabética de los términos (siete tesauros, 63%), una presentación jerárquica que muestre las relaciones e información adicional (diez tesauros, excepto Reegle por ser un glosario, 90%), una presentación gráfica que complemente a las otras (tres tesauros, Social Semantic Web, Reegle y Eurovoc, 27%) y una presentación multilingüe que facilite el uso del tesoro por parte de una gran cantidad de usuarios y aumente su visibilidad (siete tesauros, 63%). Ya hemos dicho que la norma no hace distinción entre las básicas y las complementarias.

2. En referencia a las presentaciones y diseño de los tesauros, en general estas herramientas incluyen bastantes visualizaciones de las que indica la norma, pero sería necesario que este apartado, tan básico para los usuarios, se redefiniera y se estableciera una diferencia entre las que deberían estar de forma obligatoria y las que servirían para completarlas. Por visualizaciones, el mejor tesoro en cuanto a elementos del registro simple es EUROVOC, con 7 de 8 elementos conseguidos; de la alfabética, destacan el CINDOC y el Government of Canada Core Subject, cumpliendo 3 características; en la jerárquica el balance es positivo pero sería necesario que los tesauros mejorasen en cuanto a la inclusión de algún término más en la jerarquía y la mejora de algunas características, como el sangrado. El más adaptado a la normativa es AGROVOC. La presentación gráfica es la menos desarrollada, aunque los tres tesauros que la incluyen cumplen perfectamente la función complementaria para la que sirven. El multilingüismo no aparece en todos los tesauros, y con respecto a sus características, destaca EUROVOC como única herramienta que contiene el código de dos cifras para identificar los idiomas. A nivel general, EUROVOC es el mejor tesoro en este aspecto con 19 características, algo más de la mitad de las que analizábamos.

3. En la búsqueda y navegación, los resultados generales muestran que los tesauros cumplen mejor una característica que otra, como es el caso de AGROVOC y UNESCO, que se adaptan mejor a la normativa en materia de búsqueda, o STW Economics y Social Sciences, que por el contrario, han adaptado mejor las recomendaciones del apartado de navegación. UKAT y AGROVOC incluyen ocho variables de las once que analizábamos en este apartado, por lo que son los que mejor han adaptado la norma en sus tesauros.

4. El apartado de integración es el que peor resultados ha proporcionado, aunque es necesario aclarar que hemos considerado que la norma se refiere a este aspecto teniendo en cuenta que este enlace tiene que estar en la página principal del tesoro. Sin embargo, es cierto que la mayor parte de los tesauros analizados, por no decir todos, están integrados en herramientas

documentales de distinta tipología y que realizan una función de recuperación de en otros entornos. Solo tres tesauros, UNESCO, Reegle y STW Economics están integrados con alguna aplicación, aunque cumplen muy pocas características de las dictadas por la norma.

5. A nivel general, y aunque es cierto que ha habido tesauros que han destacado por encima de otros en los distintos apartados, ninguno de ellos cumple ni siquiera el 50% de las características analizadas. Una vez aclarado este punto, es curioso que el STW Economics, habiendo cumplido más características de dos apartados, finalmente no sea el tesoro más eficiente, dejando esa afirmación para EUROVOC, que ha conseguido convertirse en la herramienta que mejor se adapta a la norma ISO 25964-1, incluyendo 26 de las 58 variables.

Por último, es necesario añadir que, debido a las carencias que hemos detectado de la norma, podemos considerar positivo que los tesauros no cumplan muchas de las características. Esto se puede interpretar como un aviso que realizan los tesauros y las instituciones para que la normativa se modifique, sobre todo, en el apartado de la visualización.

Futuras líneas de investigación

En primer lugar, es necesario que se realicen trabajos de investigación donde la muestra de tesauros sea más extensa para poder generalizar los resultados. Hay que tener en cuenta que las conclusiones que hemos extraído se basan únicamente en una reducida muestra de 11 tesauros en formato electrónico. Es una cifra muy pequeña en comparación con la cantidad de tesauros de esta tipología que hay en Internet.

También sería conveniente conocer la opinión de los usuarios en referencia a los aspectos que nosotros hemos analizado y a otros, ya que ellos son los que verdaderamente utilizan las herramientas para solucionar sus necesidades informativas. Nadie mejor que ellos pueden criticarlas y proponer soluciones de mejora.

Y por último, es necesario que se profundice más en el aspecto de la integración de los tesauros con otras herramientas y aplicaciones documentales, comprobando si verdaderamente estos instrumentos han aprovechado todas las oportunidades que ofrece la web semántica para interrelacionarse entre ellos, aprovechando así la interoperabilidad y la compartición de información. En este trabajo, nosotros hemos analizado esta integración desde el punto de vista de que se pueda acceder a otra herramienta desde la página principal del tesoro, por eso los resultados son tan negativos, pero sabemos que la mayoría de los tesauros sí que sirven como apoyo documental a otras aplicaciones.

V. BIBLIOGRAFÍA

- CARO CASTRO, Carmen (2012). Vocabularios estructurados, Web Semántica y Linked Data: oportunidades y retos para los profesionales de la documentación. En: Seminário de Arquivologia, Biblioteconomia e Ciência da Informação [en línea]: identidades, contrastes e perspectivas de interlocução (Universidade Federal Fluminense, 26-28 de setembro, 2011). Niterói (Río de Janeiro) : EdUFF, 2012, p. 127-141 [Fecha de consulta: 16 de junio de 2013]. Disponible en:
http://gredos.usal.es/jspui/bitstream/10366/121953/1/ccaro_SeminarioUFF.pdf
- CODINA, Lluís (2000). Evaluación de recursos digitales en línea: conceptos, indicadores y métodos [en línea]. *Revista Española de Documentación Científica* [en línea], 2000, v. 20, n. 1, pp. 9-44. [Fecha de consulta: 3 de mayo de 2013]. Disponible en:
<http://redc.revistas.csic.es/index.php/redc/article/view/315/479>
- DEXTRE CLARKE, Stella G. y LEI ZENG, Marcia (2012). From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modelling [en línea]. *Information Standards Quarterly (ISQ)*, 2012, v. 24, n. 1, pp. 20-26. [Fecha de consulta: 18 de mayo de 2013]. Disponible en:
http://www.niso.org/apps/group_public/download.php/8289/SP_clarke_zeng_isqv24no1.pdf
- GARCÍA MARCO, Francisco Javier (2008). Las normas de tesauros se ponen al día: vocabularios estructurados para la recuperación de información en el entorno digital. *Anuario ThinkEPI* [en línea], 2008, n. 1, pp. 57-62. [Fecha de consulta: 5 de abril de 2013]. Disponible en:
http://www.thinkepi.net/notas/2008/Garcia-Marco_normas-de-tesauros.pdf
- GARCÍA MARCO, Francisco Javier (coord.) et al (2007). Proyectos internacionales de reforma y ampliación de las normas sobre tesauros para su adaptación a los nuevos contextos de integración e interoperabilidad en el entorno digital. En: La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico [en línea]. Congreso ISKO (8º. 2007. León), 18-20 de Abril. [León] : ISKO España, 2007, pp. 389-398. [Fecha de consulta: 7 de junio de 2013]. Disponible en:
http://e-archivo.uc3m.es/bitstream/10016/4124/1/proyectos_sansegundo_ISKO_2007.pdf
- GIL URDICIÁIN, Blanca (1998a). Origen y evolución de los tesauros en España. *Revista General de Información y Documentación* [en línea], 1998a, v. 8, n. 1, pp. 63-110. [Fecha de consulta: 8 de junio de 2013]. Disponible en:
<http://revistas.ucm.es/index.php/RGID/article/view/RGID9898120063A/10796>

- GIL URDICIÁIN, Blanca (1998b). Evaluación semántica y estructural de tesauros. *Revista General de Información y Documentación* [en línea], 1998b, v. 8, n. 2, pp. 193-199. [Fecha de consulta: 6 de junio de 2013]. Disponible en: <http://revistas.ucm.es/index.php/RGID/article/view/RGID9898220193A/10755>
- GIL URDICIÁIN, Blanca (1998c). Evaluación del rendimiento de tesauros españoles en sistemas de recuperación de información. *Revista Española de Documentación Científica* [en línea], 1998c, v. 21, n. 3, pp. 286-302. Disponible en: <http://redc.revistas.csic.es/index.php/redc/article/view/357/604> [Fecha de consulta: 20 de mayo de 2013]
- ISAAC, Antoine et al (2011). Grupo Incubador de Datos Vinculados de bibliotecas: conjuntos de datos, vocabularios de valores y conjuntos de elementos de metadatos. *Informe de Grupo Incubador del W3C de 25 de Octubre 2011* [en línea]. [Fecha de consulta: 7 de junio de 2013]. Disponible en: <http://skos.um.es/Incubator/ld/XGR-ld-vocabdataset/>
- LANCASTER, Frederick Wilfrid (1995). *El control del vocabulario en la recuperación de información*. València : Universitat de València, 1995. 286 p.
- LÓPEZ-HUERTAS PÉREZ, María José (1991). *Lenguajes documentales: aproximación a la evolución histórica de un concepto*. Boletín de la ANABAD, 1991, v. 41, n.1, pp. 61-70.
- MOCHÓN BEZARES, Gonzalo y SORLI ROJO, Ángela (2010). Tesauros en acceso abierto en Internet: un análisis cuantitativo. *Revista Española de Documentación Científica* [en línea], 2010, v. 33, n. 4, pp. 643-663. [Fecha de consulta: 20 de junio de 2013]. Disponible en: <http://redc.revistas.csic.es/index.php/redc/article/view/675/750>
- PÉREZ AGÜERA, José Ramón (2004). L'automatització de tesaurus i la seva utilització en el web semàntic. *BiD: Textos universitaris de Biblioteconomia i Documentació* [en línea], 2004, n. 13. [Fecha de consulta: 20 de mayo de 2013]. Disponible en: <http://www.ub.edu/bid/13perez1.htm>
- PASTOR-SÁNCHEZ, Juan Antonio (2013). ISO-THES: Ampliando SKOS a partir de la norma de tesauros ISO-25964. *Anuario ThinkEPI* [en línea], 2013, n. 1, pp. 189-193. [Fecha de consulta: 18 de junio de 2013]. Disponible en: <http://www.thinkepi.net/iso-thes-ampliando-skos-norma-tesauros-iso-25964>
- PASTOR-SÁNCHEZ, Juan Antonio y MARTÍNEZ MÉNDEZ, José Antonio (2009). Aplicación de tesauros, taxonomías y ontologías en los sistemas de gestión de contenidos mediante tecnologías de la Web Semántica. *Ibersid: revista de sistemas de información y documentación* [en línea], 2009, v. 3, pp. 143-153. [Fecha de consulta: 26 de mayo de 2013]. Disponible en: <http://www.ibersid.eu/ojs/index.php/ibersid/article/view/3734/3495>

- PASTOR-SÁNCHEZ, Juan Antonio y SAORÍN PÉREZ, Tomás (1996). Un interfaz hipertexto para un sistema de gestión de tesauros. *Scire: Representación y organización del conocimiento* [en línea], 1996, v. 2, n. 1, pp. 51-62. [Fecha de consulta: 31 de mayo de 2013]. Disponible en: <http://www.iberid.eu/ojs/index.php/scire/article/view/1050/1032>
- PASTOR-SÁNCHEZ, Juan Antonio; MARTÍNEZ MÉNDEZ, José Antonio; y RODRÍGUEZ MUÑOZ, José Vicente (2009). Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. *Information Research* [en línea], 2009, v. 14, n. 4. [Fecha de consulta: 12 de mayo de 2013]. Disponible en: <http://informationr.net/ir/14-4/paper422.html>
- PASTOR-SÁNCHEZ, Juan Antonio; MARTÍNEZ MÉNDEZ, José Antonio y RODRÍGUEZ MUÑOZ, José Vicente (2012). Aplicación de SKOS para la interoperabilidad de vocabularios controlados en el entorno de Linked Open Data. *El Profesional de la Información* [en línea], 2012, v. 21, n. 3, pp. 245-253. [Fecha de consulta: 5 de mayo de 2013]. Disponible en: <http://eprints.rclis.org/17066/1/245-253.pdf>
- SÁNCHEZ-CUADRADO, Sonia; COLMENERO-RUIZ, María-Jesús y MOREIRO, José Antonio (2012). Tesauros: estándares y recomendaciones. *El Profesional de la Información* [en línea], 2012, v. 21, n. 3, pp. 229-235. [Fecha de consulta: 31 de mayo de 2013]. Disponible en: <http://eprints.rclis.org/17600/1/229-235.pdf>
- TESAUROS: estándares y recomendaciones [en línea]. [Fecha de consulta: 5 de mayo de 2013]. Disponible en: <http://web.usal.es/~alar/Bibweb/Temario/Tesauro.PDF>
- VALLE BRACERO, Antonio et al (2000). Tesauros en HTML. Un modelo de diseño y estructura para su consulta en la malla mundial (WWW). *Revista Española de Documentación Científica* [en línea], 2000, v. 23, n. 2, pp. 159-178. Disponible en: [Fecha de consulta: 19 de mayo de 2013] <http://redc.revistas.csic.es/index.php/redc/article/view/321/488>
- WILLIAMSON, Nancy J (2007). Knowledge structures and the Internet: progress and prospects [en línea]. *Cataloging & Classification Quarterly (CCQ)*, 2007, v. 44, n. 3-4, pp. 329-342. [Fecha de consulta: 3 de junio de 2013]. Disponible en: https://ualearn.blackboard.com/bbcswebdav/pid-603350-dt-content-rid-2678521_1/courses/18998.201310/readings/williamson.pdf

Normativa

DIRECTRICES para el establecimiento y desarrollo de tesauros monolingües [en línea]. UNE 50106:1990. En: *Documentación. T. 2, Normas fundamentales*. Madrid: AENOR, 1997; pp. 77-126. Disponible en: http://elmemorioso.cl/une50106_90.pdf [Fecha de consulta: 8 de marzo de 2013]

DIRECTRICES para la creación y desarrollo de tesauros multilingües. UNE 50125:1997. En: *Documentación. T. 2, Normas fundamentales*. Madrid: AENOR, 1997; pp. 320-395. [Fecha de consulta: 8 de marzo de 2013]

GUIDELINES for the establishment and development of monolingual thesauri. Norme International ISO 2788:1986. [S.l.]: ISO, 1986.

GUIDELINES for the establishment and development of multilingual thesauri. Norme international ISO 5964:1985. [S.l.]: ISO, 1985.

THESAURI and interoperability with other vocabularies — Part 1: Thesauri for information retrieval. Norme international ISO 25964-1:2011. [S.l.]: ISO, 2011, 143 p.

Recursos web

AGROVOC Agricultural Thesaurus [en línea]. [S.l.]: FAO, 2012. [Fecha de consulta: 5 de mayo de 2013]. Disponible en: <http://aims.fao.org/standards/agrovoc/about>

EUROVOC. Tesauro multilingüe de la Unión Europea [en línea]. [S.l.] : Oficina de Publicaciones de la Unión Europea, 200-. Disponible en: <http://eurovoc.europa.eu/drupal/?q=es> [Fecha de consulta: 3 de mayo de 2013]

GEMET (General Multilingual Environmental Thesaurus) [en línea]. [S.l.]: GEMET, 2010. [Fecha de consulta: 1 de mayo de 2013]. Disponible en: <http://www.eionet.europa.eu/gemet>

GOVERNMENT of Canada Core Subject Thesaurus [en línea]. [S.l.]: Government of Canada, 2013. [Fecha de consulta: 29 de abril de 2013]. Disponible en: <http://en.thesaurus.gc.ca/default.asp?Lang=En&n=E5807AB0-1>

INSTITUTO de Estudios Documentales sobre Ciencia y Tecnología (IEDCYT) (200-) [en línea]. [S.l.]: [s.n.], [200-]. [Fecha de consulta: 11 de mayo de 2013]. Disponible en: http://thes.cindoc.csic.es/index_esp.php

MOCHÓN BEZARES, Gonzalo y SORLI ROJO, Ángela (2005). Tesoro de Biblioteconomía y Documentación [en línea]. [S.l.]: [s.n.], 2005. [Fecha de consulta: 6 de mayo de 2013]. Disponible en:

http://thes.cindoc.csic.es/index_BIBLIO_esp.php?thes=BIBLIO

REEGLE - Clean Energy and Climate Change Thesaurus (200-) [en línea]. Viena : Renewable Energy & Energy Efficiency Partnership, Paris: Renewable Energy Policy Network for the 21st Century, [200-]. [Fecha de consulta: 2 de mayo de 2013]. Disponible en:

<http://data.reegle.info/thesaurus/guide>

SEMANTIC Web Deployment Working Group (2013). *"Simple Knowledge Organization System (SKOS): SKOS/Dataset"* [en línea]. [S.l.] : [s.n.], 2013. [Fecha de consulta: 29 de abril de 2013]. Disponible en:

<http://www.w3.org/2001/sw/wiki/SKOS/Datasets>

SOCIAL Semantic Web Thesaurus [en línea]. Viena: Semantic Web Company GMBH, 2013. [Fecha de consulta: 4 de mayo de 2013]. Disponible en: <http://vocabulary.semantic-web.at/semweb.html>

STW Thesaurus for Economics [en línea]. Kiel (Alemania) : ZBW – German National Library of Economics Leibniz Information Centre for Economics, 2012. [Fecha de consulta: 2 de mayo de 2013]. Disponible en: <http://zbw.eu/stw/versions/latest/about>

THESAURUS for the Social Sciences [en línea]. Mannheim (Alemania): GESIS – Leibniz Institute for the Social Sciences, 2013. [Fecha de consulta: 4 de mayo de 2013]. Disponible en: <http://www.gesis.org/en/services/research/thesauri-und-klassifikationen/social-science-thesaurus/>

UK Archival Thesaurus (UKAT) (2004) [en línea]. Londres : UKAT, 2004. [Fecha de consulta: 6 de mayo de 2013]. Disponible en: <http://www.ukat.org.uk/>

UNESCO Thesaurus [en línea]. París : UNESCO, 2010. [Fecha de consulta: 7 de mayo de 2013] Disponible en: <http://databases.unesco.org/thessp/>

VI. ANEXOS

Anexo I – Plantilla para la evaluación de los tesauros

a) Registro simple

1. ¿El tesauro permite ver registros simples?

a) Sí

b) No

2. Indique cuáles de los siguientes elementos aparecen en el registro simple:

a) Código del concepto

b) Notas (de alcance, definición, historia, etc.)

c) Términos no preferentes

d) Término cabecera

e) Término general

f) Término específico

g) Término relacionado

j) Microtesauro

b) Visualización alfabética

3. ¿El tesauro tiene visualización alfabética?

a) Sí

b) No

4. Indique cuáles de estas características aparecen la presentación alfabética:

a) Los términos (preferentes y no preferentes) están ordenados en una secuencia alfabética única

b) Los términos preferentes se distinguen de los no preferentes tipográficamente

c) Se incluye el registro completo de cada término como si fueran registros simples

d) Se muestran los términos genéricos y específicos más inmediatos

e) Se utiliza la sangría para los términos subordinados

f) Se presentan los términos (preferentes y no preferentes) en un índice permutado

c) Visualización jerárquica/sistemática

5. ¿El tesauro tiene visualización sistemática?

a) Sí

b) No

6. En el primer nivel de esta visualización los términos están organizados por:

- a) Temas
- b) Facetas

7. ¿Existe una lista de términos generales que no se asocian con ninguna categoría en concreto, sino que son válidos para todas?

- a) Sí
- b) No

8. ¿Aparece el término general seguido del resto de los términos con distintos niveles de sangría?

- a) Sí
- b) No

9. ¿Los términos están organizados en facetas dentro de los campos temáticos?

- a) Sí
- b) No

10. ¿Se incluye información adicional de cada término?

- a) Sí
- b) No

11. Indique qué información adicional se añade:

- a) Código del concepto
- b) Nota de alcance
- c) Términos no preferentes
- d) Término relacionado
- e) Equivalencias interlingüísticas

d) Visualización gráfica

12. ¿El tesauro tiene visualización gráfica?

- a) Sí
- b) No

e) Presentaciones multilingües

13. ¿Se puede consultar el tesauro en varios idiomas?

- a) Sí
- b) No

14. ¿Cada idioma se identifica con un código de dos dígitos siguiendo la norma ISO 639-1?

- a) Sí
- b) No

15. ¿La secuencia alfabética es distinta para cada idioma?

- a) Sí
- b) No

16. ¿Cada término preferente tiene su equivalencia interlingüística?

- a) Sí
- b) No

17. ¿Se permite la visualización simultánea en dos idiomas?

- a) Sí
- b) No

18. ¿Las equivalencias interlingüísticas se muestran en una tabla de correspondencia independiente?

- a) Sí
- b) No

19. ¿Se permite cambiar el idioma de la pantalla?

- a) Sí
- b) No

f) Exportación

20. ¿Existe una opción que permita exportar el tesoro en un formato estándar, ya sea todo o una parte?

- a) Sí
- b) No

g) Opciones de búsqueda y navegación en el tesoro

21. ¿Es posible buscar por términos que contienen cualquier palabra?

- a) Sí
- b) No

22. ¿Es posible buscar por términos que contienen una combinación de palabras?

- a) Sí
- b) No

23. ¿Es posible buscar por términos que contienen una combinación de caracteres?

- a) Sí
- b) No

24. ¿Es posible buscar por términos y conceptos utilizando el código que los identifica?

- a) Sí
- b) No

25. ¿Se facilita la posibilidad de mostrar términos equivalentes en el mismo idioma?

- a) Sí
- b) No

26. ¿Se facilita la posibilidad de mostrar términos equivalentes en distinto idioma?

- a) Sí
- b) No

27. ¿Es posible moverse desde el registro de un concepto hasta sus términos más genéricos?

- a) Sí
- b) No

28. ¿Es posible moverse desde el registro de un concepto hasta sus términos más específicos?

- a) Sí
- b) No

29. ¿Se pueden contraer y expandir los niveles de las visualizaciones jerárquicas/sistemáticas?

- a) Sí
- b) No

30. ¿Se permite cambiar de un tipo de visualización a otra?

- a) Sí
- b) No

31. ¿Se pueden ver dos visualizaciones de forma simultánea?

- a) Sí
- b) No

h) Integración con las aplicaciones

32. ¿El tesauro está integrado con las funciones de búsqueda de una base de datos documental?

- a) Sí
- b) No

33. ¿El tesauro da la opción de seleccionar un término que se utiliza en una búsqueda?
- a) Sí
 - b) No
34. ¿El tesauro da la opción de seleccionar más de un término que se utiliza en una búsqueda?
- a) Sí
 - b) No
35. ¿El tesauro tiene la opción de construir una estrategia de búsqueda copiando términos desde el tesauro?
- a) Sí
 - b) No
36. ¿El tesauro tiene la opción de “explode”?
- a) Sí
 - b) No
37. En la búsqueda, ¿se sustituye automáticamente un término no preferente por uno preferente?
- a) Sí
 - b) No
38. En el proceso de refinamiento de la búsqueda, ¿el tesauro proporciona el número de publicaciones que tiene cada subconjunto de la colección?
- a) Sí
 - b) No
39. Cuando un usuario introduce un término similar a uno o más términos del tesauro, ¿se responde con “¿Quiso decir...?”, de manera que se ofrece al usuario coincidencias cercanas de términos que se encuentran en el tesauro?
- a) Sí
 - b) No
40. Cuando un usuario introduce un término similar a uno o más términos del tesauro, ¿éste responde recuperando los resultados de todos los términos y agrupándolos?
- a) Sí
 - b) No
41. ¿El usuario puede ampliar o delimitar su búsqueda?
- a) Sí
 - b) No

42. Si la búsqueda da pocos resultados, ¿se puede expandir automáticamente a otros términos? (generales, específicos, relacionados...)

- a) Sí
- b) No

43. Cuando un término de búsqueda se transforma, sustituye, traduce o reemplaza por otro, ¿se informa al usuario de este hecho?

- a) Sí
- b) No